



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 181755**

**TO: Celine Qian**  
**Location: rem/2A64/2C70**  
**Art Unit: 1636**  
*March 14, 2006*  
**Case Serial Number: 09/515363**

**From: P. Sheppard**  
**Location: Remsen Building**  
**Phone: (571) 272-2529**

**sheppard@uspto.gov**

### **Search Notes**

**This Page Blank (uspto)**

FOR OFFICIAL USE ONLY

Scientific and Technical Information Center

CRFE

SEARCH REQUEST FORM

Requester's Full Name: Celine Qian Examiner #: 78710 Date: 3/8/06  
Art Unit: 1636 Phone Number: 2- 0777 Serial Number: 09/515363  
Location (Bldg/Room#): 2464 (Mailbox #): 2070 Results Format Preferred (circle): PAPER DISK  
\*\*\*\*\*

To ensure an efficient and quality search, please attach a copy of the cover sheet, claims, and abstract or fill out the following:

Title of Invention: Melanoma Diff. Assoc. Gene, Promoter & uses thereof.  
Inventors (please provide full names): Fish, Paul et al.

Earliest Priority Date: 2/29/2000

Search Topic:

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known.

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

MEY Please search SEQ ID NO: 1  
and a nucleic acid encoding SEQ ID NO: 2  
Please search above sequences in both commercial  
database & interference database.  
1 NA 3365  
✓ 2 AA 1025

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2006, 18:20:20 ; Search time 16315 Seconds  
(without alignments)  
11724.071 Million cell updates/sec

Title: US-09-515-363C-1  
Perfect score: 3365  
Sequence: 1 gcgcgcgcgcctgagagccc.....aaatgattgttactctg 3365

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_sts:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_hcg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3365	100.0	3380	6	CQ776470	CQ776470 Sequence
2	3365	100.0	3380	6	CS108225	CS108225 Sequence
3	3365	100.0	3380	6	CS113027	CS113027 Sequence
4	3365	100.0	3380	6	AF095844	AF095844 Homo sapi
5	3361.8	99.9	3434	6	CS033863	CS033863 Sequence
6	3361.8	99.9	3434	6	CS036676	CS036676 Sequence
7	3361.8	99.9	3434	6	CS042815	CS042815 Sequence
8	3361.8	99.9	3434	6	CS045628	CS045628 Sequence
9	3338.8	99.2	3372	6	AX300832	AX300832 Sequence
10	3338.8	99.2	3373	8	AY017378	AY017378 Homo sapi
11	2476	73.6	2540	8	BC078180	BC078180 Homo sapi
12	2202.4	65.5	3771	6	CQ777379	CQ777379 Sequence
13	2202.4	65.5	3771	9	AF374384	AF374384 Mus muscu
14	2135.4	63.5	3078	9	AY075132	AY075132 Mus muscu
15	1995.4	59.3	3635	9	BC080200	BC080200 Mus muscu
16	1766	52.5	1776	6	AX714144	AX714144 Sequence
17	1766	52.5	1776	8	AK056293	AK056293 Homo sapi
18	1578.8	46.9	2028	6	CQ729104	CQ729104 Sequence

19	1435.6	42.7	1443	6	AX300841	AX300841 Sequence
20	1402	41.7	2468	9	BC025508	BC025508 Mus muscu
21	1278.8	38.0	1392	6	AR379503	AR379503 Sequence
22	1254.6	37.3	1284	6	AX300838	AX300838 Sequence
23	1193.4	35.5	1204	8	BC007966	BC007966 Homo sapi
24	1186	35.2	1250	9	BC004031	BC004031 Mus muscu
25	803	23.9	1258	6	AX098232	AX098232 Sequence
26	803	23.9	1270	6	AX098236	AX098236 Sequence
27	797	23.7	1630	8	BC046208	BC046208 Homo sapi
28	663	19.7	3692	6	AX098234	AX098234 Sequence
29	663	19.7	3704	6	AX098238	AX098238 Sequence
30	623.2	18.5	95417	8	AC010876	AC010876 Homo sapi
31	507.2	15.1	595	6	CQ100724	CQ100724 Sequence
32	507.2	15.1	595	6	CQ139716	CQ139716 Sequence
33	507.2	15.1	595	6	CQ261044	CQ261044 Sequence
34	507.2	15.1	595	6	CQ335189	CQ335189 Sequence
35	455.2	13.5	499	6	CQ113866	CQ113866 Sequence
36	455.2	13.5	499	6	CQ152744	CQ152744 Sequence
37	455.2	13.5	499	6	CQ273644	CQ273644 Sequence
38	455.2	13.5	499	6	CQ347943	CQ347943 Sequence
39	440.8	13.1	444	6	CQ729107	CQ729107 Sequence
40	384	11.4	923	9	BC019605	BC019605 Mus muscu
41	376.4	11.2	458	6	BD269355	BD269355 Gene show
42	371.2	11.0	392	6	AX071769	AX071769 Sequence
43	365.8	10.9	374	6	BD071774	BD071774 Secreted
44	356.8	10.6	387	6	AX071783	AX071783 Sequence
45	336	10.0	3458	5	BC073528	BC073528 Xenopus 1

ALIGNMENTS

RESULT 1  
LOCUS CQ776470 3380 bp DNA linear PAT 11-MAR-2004  
DEFINITION Sequence 156 from Patent EP1394274.  
ACCESSION CQ776470  
VERSION CQ776470.1 GI:45379860  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuhara,K. Methods of testing for bronchial asthma or chronic obstructive pulmonary disease Patent: EP 1394274-A 156 03-MAR-2004;  
JOURNAL Genex Research, Inc. (JP)  
FEATURES  
source 1. 3380  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 3365; DB 6; Length 3380;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGGCGCGGCTGAGAGCCCTGTGACAACTGTCATTGTCAGGACAGAGCGGTAGAC 60  
Db 1 GCGGCGCGGCTGAGAGCCCTGTGACAACTGTCATTGTCAGGACAGAGCGGTAGAC 60  
QY CCTGCTTCTTAAGTGGGACGGACAGCGGACGACATTTCACCTGTCCCGACAGCAA 120  
Db CCTGCTTCTTAAGTGGGACGGACAGCGGACGACATTTCACCTGTCCCGACAGCAA 120  
QY CAGCACCATCTGCTTGGAGAACCTCTCCCTTCTGAGAAAGAAAGATGCGAATGCG 180  
Db CAGCACCATCTGCTTGGAGAACCTCTCCCTTCTGAGAAAGAAAGATGCGAATGCG 180

QY 181 TATTCACAGACGAGATTTCCGCTATCTCATCTCGTCTCAGGCCACAGGTGAATG 240  
|||||  
Db 181 TATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTCTCAGGCCACAGGTGAATG 240  
QY 241 TACATCCAGGTGGAGCCGTGTGCTGACTAAGCTTTCTGCTCTCAGAGGTGAAGAG 300  
|||||  
Db 241 TACATCCAGGTGGAGCCGTGTGCTGACTAAGCTTTCTGCTCTCAGAGGTGAAGAG 300  
QY 301 CAGATTGAGAGCAGTCCGCACTCCGGGAACATGACAGGCAGTTGAATGCTGCTGAGC 360  
|||||  
Db 301 CAGATTGAGAGCAGTCCGCACTCCGGGAACATGACAGGCAGTTGAATGCTGCTGAGC 360  
QY 361 ACCTTGAGAGAGAGTCTGCACTTGGTTGGAATCGGGAATTCGTGAGGCCCTCGG 420  
|||||  
Db 361 ACCTTGAGAGAGAGTCTGCACTTGGTTGGAATTCGTGAGGCCCTCGG 420  
QY 421 AGAACCCGCGACCCCTCTGGCCGCGCTACATGAACCCCTGAGCTCACGGACTTGCCCTCT 480  
|||||  
Db 421 AGAACCCGCGACCCCTCTGGCCGCGCTACATGAACCCCTGAGCTCACGGACTTGCCCTCT 480  
QY 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTCAAGCCCACT 540  
|||||  
Db 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTCAAGCCCACT 540  
QY 541 CTGTGACAGAGCTTCTAGTTAGAGAGCTCTTGATTAAGTGCATGAGAGGAATCTGTTG 600  
|||||  
Db 541 CTGTGACAGAGCTTCTAGTTAGAGAGCTCTTGATTAAGTGCATGAGAGGAATCTGTTG 600  
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAACCAATGGAATGATCAGTGTGA 660  
|||||  
Db 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAACCAATGGAATGATCAGTGTGA 660  
QY 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAATGCTGCTCATTTCTGAATGTT 720  
|||||  
Db 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAATGCTGCTCATTTCTGAATGTT 720  
QY 721 CTTGCTCAAAACAGAAACCAATGAATCTTGTCCAAGATTAAACAGGCTCTGATTGCTCAGAA 780  
|||||  
Db 721 CTTGCTCAAAACAGAAACCAATGAATCTTGTCCAAGATTAAACAGGCTCTGATTGCTCAGAA 780  
QY 781 AGCAATGCAGAGATTGAGAAATTTATCAGAAGTTGATGTCCTCAAGTGAAGAGCAACTT 840  
|||||  
Db 781 AGCAATGCAGAGATTGAGAAATTTATCAGAAGTTGATGTCCTCAAGTGAAGAGCAACTT 840  
QY 841 CTTTCAACCAAGTTCAAGCAATCTGGAAGAGAGTCTGGGCATGAGAAATTAAGTCA 900  
|||||  
Db 841 CTTTCAACCAAGTTCAAGCAATCTGGAAGAGAGTCTGGGCATGAGAAATTAAGTCA 900  
QY 901 TCAGATCATCTTTGACAGATTCTCTGTAAGTTTCAGAATCAGACACAAGTTGGCAGAA 960  
|||||  
Db 901 TCAGATCATCTTTGACAGATTCTCTGTAAGTTTCAGAATCAGACACAAGTTGGCAGAA 960  
QY 961 GGAAGTGTCAAGTCTTATGATGAAAAGTCTTGAACATTAACAGCAACATGGGCAGTATCA 1020  
|||||  
Db 961 GGAAGTGTCAAGTCTTATGATGAAAAGTCTTGAACATTAACAGCAACATGGGCAGTATCA 1020  
QY 1021 GGCACCATGGAGAGTATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCA 1080  
|||||  
Db 1021 GGCACCATGGAGAGTATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCA 1080  
QY 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCAGCTTGAAGGGAAGAAAT 1140  
|||||  
Db 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCAGCTTGAAGGGAAGAAAT 1140  
QY 1141 ATCATCATCTGCTCCCTTACAGGAGTGAAGAAACAGAGTGGCTTTTACATTGCCAAG 1200  
|||||  
Db 1141 ATCATCATCTGCTCCCTTACAGGAGTGAAGAAACAGAGTGGCTTTTACATTGCCAAG 1200  
QY 1201 GATCACTTAGACAGAGAAAAAGCATCTGAGCCCTGGAAGAAATTAAGTTCTTGTCAAT 1260  
|||||  
Db 1201 GATCACTTAGACAGAGAAAAAGCATCTGAGCCCTGGAAGAAATTAAGTTCTTGTCAAT 1260  
QY 1261 AAGTACTGCTAAGTTGAACAGCTCTTCCGCAAGAGATTCCAACCAATTTTGAAGAAATGG 1320

Db 1261 AAGTACTGCTAAGTTGAACAGCTCTTCCGCAAGAGATTCCAACCAATTTTGAAGAAATGG 1320  
QY 1321 TATCGTTTATTTGATTAAAGTGTGATACCCCAACTGAAAAATATCATTTCCAGAAATTGTC 1380  
|||||  
Db 1321 TATCGTTTATTTGATTAAAGTGTGATACCCCAACTGAAAAATATCATTTCCAGAAATTGTC 1380  
QY 1381 AAGTCTGTGATATTATATCAGTACAGCTCAAAATCCTGAAAAATCCCTCTTAAACTTG 1440  
|||||  
Db 1381 AAGTCTGTGATATTATATCAGTACAGCTCAAAATCCTGAAAAATCCCTCTTAAACTTG 1440  
QY 1441 GAAATGAGAGAGATGCTGTTCAATTGTACAGACTTTTCCCTCATTATCATGTAGTAA 1500  
|||||  
Db 1441 GAAATGAGAGAGATGCTGTTCAATTGTACAGACTTTTCCCTCATTATCATGTAGTAA 1500  
QY 1501 TGTCAATCAACCAACAAAGAGAGTGTATTAATTAATGAGGCAATTTGATGCAAG 1560  
|||||  
Db 1501 TGTCAATCAACCAACAAAGAGAGTGTATTAATTAATGAGGCAATTTGATGCAAG 1560  
QY 1561 AAGTTGAAAAAATATAGACTCAAGAAAGAAAAAACAACAGTGAATCCCTTCTCAGATA 1620  
|||||  
Db 1561 AAGTTGAAAAAATATAGACTCAAGAAAGAAAAAACAACAGTGAATCCCTTCTCAGATA 1620  
QY 1621 CTGGAGCTAACAGCTTCACTGCTGTGTTGAGAGGGGCCAGAACCAAGCCAAAGCTGAAGAA 1680  
|||||  
Db 1621 CTGGAGCTAACAGCTTCACTGCTGTGTTGAGAGGGGCCAGAACCAAGCCAAAGCTGAAGAA 1680  
QY 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTTACTATTAATACTGTTAAAGAAAAAC 1740  
|||||  
Db 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTTACTATTAATACTGTTAAAGAAAAAC 1740  
QY 1741 CTGATCAACTGAAAAAACCATAATACAGAGCCATGCAAGAAAGTTGGCATTTGCAGATGCA 1800  
|||||  
Db 1741 CTGATCAACTGAAAAAACCATAATACAGAGCCATGCAAGAAAGTTGGCATTTGCAGATGCA 1800  
QY 1801 ACCAGAGAGATCCATTAAAGAGAACTTTAGAAATTAATGACAAAGATTTCAACTTAT 1860  
|||||  
Db 1801 ACCAGAGAGATCCATTAAAGAGAACTTTAGAAATTAATGACAAAGATTTCAACTTAT 1860  
QY 1861 TGTCAATGAGTCCAATGTCAAGTTTGAAGTCAACCTCATGAAACAATGGGCCATTCAA 1920  
|||||  
Db 1861 TGTCAATGAGTCCAATGTCAAGTTTGAAGTCAACCTCATGAAACAATGGGCCATTCAA 1920  
QY 1921 ATGAAAAAAGCTGCAAAAAAGAAATCGCAAGAAAGTGTGTCAGAAACTTTG 1980  
|||||  
Db 1921 ATGAAAAAAGCTGCAAAAAAGAAATCGCAAGAAAGTGTGTCAGAAACTTTG 1980  
QY 1981 AGGAGTACATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTAACT 2040  
|||||  
Db 1981 AGGAGTACATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTAACT 2040  
QY 2041 CATCTTGAAACTTTCTATTAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAATGAT 2100  
|||||  
Db 2041 CATCTTGAAACTTTCTATTAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAATGAT 2100  
QY 2101 AGTGATGAGGTGCTGATGATGATGATTTGTGATGTTGATGATGAGATGAGATTTAAAG 2160  
|||||  
Db 2101 AGTGATGAGGTGCTGATGATGATGATTTGTGATGTTGATGATGAGATGAGATTTAAAG 2160  
QY 2161 AAACCTTTGAAACTGATGAAGAGATGATTTCTCATGACTTTATTTTGAAGAACAAT 2220  
|||||  
Db 2161 AAACCTTTGAAACTGATGAAGAGATGATTTCTCATGACTTTATTTTGAAGAACAAT 2220  
QY 2221 AAAATGTTGAAAAAGGCTGGCTGAAAAACCAAGATATGAAAAAGCTGACCAATTA 2280  
|||||  
Db 2221 AAAATGTTGAAAAAGGCTGGCTGAAAAACCAAGATATGAAAAAGCTGACCAATTA 2280  
QY 2281 AGAATATACATTAATGAGCAATATTAAGACTGAGGAAATCAGACAGAGGAATTAATCTTT 2340  
|||||  
Db 2281 AGAATATACATTAATGAGCAATATTAAGACTGAGGAAATCAGACAGAGGAATTAATCTTT 2340  
QY 2341 ACAAAAAACAGACAGAGTGCATATGCGCTTTCCAGTGAATTAAGTGAAGAAATTTT 2400  
|||||

Db 2341 ACAAAAACACGACAGAGTGCATATGCGCTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400  
QY 2401 GCTGAAGTAGAGTCAAAAGCCCCACCATCTGATTGAGCTGGACACAGCAGTGAATTCAA 2460  
Db 2401 GCTGAAGTAGAGTCAAAAGCCCCACCATCTGATTGAGCTGGACACAGCAGTGAATTCAA 2460  
QY 2461 CCCATGACACAGAAATGACAAAAAGAGTCAATTAATTTCCGACCTGGAAAAATCAAT 2520  
Db 2461 CCCATGACACAGAAATGACAAAAAGAGTCAATTAATTTCCGACCTGGAAAAATCAAT 2520  
QY 2521 CTGCTTATCGCTACACAGTGGCAGAGAAGGCTGGATATTAAGAATGTAACATTGTT 2580  
Db 2521 CTGCTTATCGCTACACAGTGGCAGAGAAGGCTGGATATTAAGAATGTAACATTGTT 2580  
QY 2581 ATCCGTTATGCTCTCGTCAACCAATGAATAGCCATGCTCCAGGCCGCTGTCAGCCAGA 2640  
Db 2581 ATCCGTTATGCTCTCGTCAACCAATGAATAGCCATGCTCCAGGCCGCTGTCAGCCAGA 2640  
QY 2641 GCTGATGAGAGCACCCTACGTCCTGCTGCTCAGACAGTGGTTCAGAGATTTCGAACATGAG 2700  
Db 2641 GCTGATGAGAGCACCCTACGTCCTGCTGCTCAGACAGTGGTTCAGAGATTTCGAACATGAG 2700  
QY 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAGCTATACATTGTCCTCAAAATATG 2760  
Db 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAGCTATACATTGTCCTCAAAATATG 2760  
QY 2761 AAACGAGAGAGATATGCTCATAGATTTTGAATTAACAGATGCAAGATGTAATGAGAAAG 2820  
Db 2761 AAACGAGAGAGATATGCTCATAGATTTTGAATTAACAGATGCAAGATGTAATGAGAAAG 2820  
QY 2821 AAAATGAAAAACCAAGAGAAATATTTGCCAAGCATTAACAAGATAACCCATCACTAATACT 2880  
Db 2821 AAAATGAAAAACCAAGAGAAATATTTGCCAAGCATTAACAAGATAACCCATCACTAATACT 2880  
QY 2881 TTCCTTTGCAAAAACCTGCACTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAG 2940  
Db 2881 TTCCTTTGCAAAAACCTGCACTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAG 2940  
QY 2941 AAAATGCAATCAGTCAATATGACCCCAAGAACTTACATTTGTAAGAGAAAAAC 3000  
Db 2941 AAAATGCAATCAGTCAATATGACCCCAAGAACTTACATTTGTAAGAGAAAAAC 3000  
QY 3001 AAAGCACTGCAAAAAGAGTGGCCGACTATCAAAATTAATGTGAATCATCTGCAAAATGT 3060  
Db 3001 AAAGCACTGCAAAAAGAGTGGCCGACTATCAAAATTAATGTGAATCATCTGCAAAATGT 3060  
QY 3061 GGGCAGGCTTGGGGAACAATGATGTGCAACAAGGCTTAGATTGCTCTCAAAATA 3120  
Db 3061 GGGCAGGCTTGGGGAACAATGATGTGCAACAAGGCTTAGATTGCTCTCAAAATA 3120  
QY 3121 AGGAATTTTGTAGTGTTCAAAAATTAATTAACAAGAAAAAATAACATAACAAAAAGTGGA 3180  
Db 3121 AGGAATTTTGTAGTGTTCAAAAATTAATTAACAAGAAAAAATAACATAACAAAAAGTGGA 3180  
QY 3181 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3240  
Db 3181 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3240  
QY 3241 GATTAGCACTGATGTAAGATTTCTTTAAATATCAATCAAGTTAAACATTTAATATGATTA 3300  
Db 3241 GATTAGCACTGATGTAAGATTTCTTTAAATATCAATCAAGTTAAACATTTAATATGATTA 3300  
QY 3301 TGATTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3360  
Db 3301 TGATTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3360  
QY 3361 CTCTG 3365  
Db 3361 CTCTG 3365

LOCUS CS108225 3380 bp DNA linear PAT 22-JUN-2005  
DEFINITION Sequence 233 from Patent WO2005051988.  
ACCESSION CS108225  
VERSION CS108225.1 GI:68147706  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 Abbas,A., Bodary,S., Clark,H., Schoenfeld,J., Williams,P.M., Wood,W.I. and Wu,T.D. Compositions and methods for the treatment of systemic lupus erythematosus Patent: WO 2005051988-A 233 09-JUN-2005; Genentech, Inc. (US)  
FEATURES  
source 1. .3380 /organism="Homo sapiens" /mol\_type="unassigned DNA" /db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 3365; DB 6; Length 3380;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGGGCCGGCCTGAGAGCCCTGTGACAACTCGTATTGTGAGGACAGAGCGTAGAC 60  
Db 1 GGGGGCCGGCCTGAGAGCCCTGTGACAACTCGTATTGTGAGGACAGAGCGTAGAC 60  
QY 61 CCTGCTTCTTAAGTGGGACGGGACAGCGGACGACATTTCACCTGTCCCGCAGACAA 120  
Db 61 CCTGCTTCTTAAGTGGGACGGGACAGCGGACGACATTTCACCTGTCCCGCAGACAA 120  
QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGCGAATGG 180  
Db 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGCGAATGG 180  
QY 181 TATTCACAGACGAAATTCGCTATCTCATCTCGTCTTCAAGGCGCAGGCTGAAAATG 240  
Db 181 TATTCACAGACGAAATTCGCTATCTCATCTCGTCTTCAAGGCGCAGGCTGAAAATG 240  
QY 241 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCTGCTGAGAGGTGAAGAG 300  
Db 241 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCTGCTGAGAGGTGAAGAG 300  
QY 241 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCTGCTGAGAGGTGAAGAG 300  
Db 241 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCTGCTGAGAGGTGAAGAG 300  
QY 301 CAGATTCAAGAGACAGTCCCACTCCGGGAACATGCAAGGCACTGTAATGCTGAGC 360  
Db 301 CAGATTCAAGAGACAGTCCCACTCCGGGAACATGCAAGGCACTGTAATGCTGAGC 360  
QY 361 ACCTTGAGAAAGGAGTCTGGCACCTTGGTTGACTCGGGAATTCGTGAGAGGCCCTCCG 420  
Db 361 ACCTTGAGAAAGGAGTCTGGCACCTTGGTTGACTCGGGAATTCGTGAGAGGCCCTCCG 420  
QY 421 AGAACCCGAGCCCTCTGGCCCGCCGCTACATGAACCTGAGCTCAGGACTTGCCTCT 480  
Db 421 AGAACCCGAGCCCTCTGGCCCGCCGCTACATGAACCTGAGCTCAGGACTTGCCTCT 480  
QY 481 CCATCGTTTGAAGACGCTCATGATGAATATCTCAACTGCTGAACCTCTTCAAGCCACT 540  
Db 481 CCATCGTTTGAAGACGCTCATGATGAATATCTCAACTGCTGAACCTCTTCAAGCCACT 540  
QY 541 CTGGTGAACAAGCTTCTAGTTAGAGAGCTCTTGATTAAGTGCATGAGAGGAAGTGTG 600  
Db 541 CTGGTGAACAAGCTTCTAGTTAGAGAGCTCTTGATTAAGTGCATGAGAGGAAGTGTG 600  
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGCTGAGAAAAAATGAAATGAATCAGGTGTA 660  
Db 601 ACAATTGAAGACAGAAACCGGATTGCTGCTGAGAAAAAATGAAATGAATCAGGTGTA 660  
QY 661 AGAGAGCTACTAAAAAGATTGTGACAGAAAGAAAACTGTTCTCTGCAATTTCTGAATGTT 720

Db 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTGTGCAATTTCTGAATGTT 720  
Qy 721 CTTGCGTCAAAACAGAAACAATGAACTTGTGTCCAAAGTTAACAGGCTCTGATGTCTCAGAA 780  
Db 721 CTTGCGTCAAAACAGAAACAATGAACTTGTGTCCAAAGTTAACAGGCTCTGATGTCTCAGAA 780  
Qy 781 AGCAATGCAGAGATTGAGATTATCAACAAGTTGATGGTCCCTCAAGTGAAGAGCAACTT 840  
Db 781 AGCAATGCAGAGATTGAGATTATCAACAAGTTGATGGTCCCTCAAGTGAAGAGCAACTT 840  
Qy 841 CTTTCAACCAACAGGTTCAAGCCAATCTGAGAGAGAGGCTCTGGGGCATGGAGATACTCA 900  
Db 841 CTTTCAACCAACAGGTTCAAGCCAATCTGAGAGAGAGGCTCTGGGGCATGGAGATACTCA 900  
Qy 901 TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCAAGAAATCAGACACAACTTTGGCAGAA 960  
Db 901 TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCAAGAAATCAGACACAACTTTGGCAGAA 960  
Qy 961 GGAAGTGTCAAGCTGCTTAGATGAAAAGTCTTGACATTAACAGCAACATGGGCACTGATTCA 1020  
Db 961 GGAAGTGTCAAGCTGCTTAGATGAAAAGTCTTGACATTAACAGCAACATGGGCACTGATTCA 1020  
Qy 1021 GGCAACCATGGGAAGTATTCAGATGAAGAGAAATGTGCAGCAAGAGCATCCCCGGAGCCA 1080  
Db 1021 GGCAACCATGGGAAGTATTCAGATGAAGAGAAATGTGCAGCAAGAGCATCCCCGGAGCCA 1080  
Qy 1081 GAACTCCAGCTCAGGCGCTTACCAATGGAAGTTGCCCAAGCCAGCCCTTGGAAAGGAAGAT 1140  
Db 1081 GAACTCCAGCTCAGGCGCTTACCAATGGAAGTTGCCCAAGCCAGCCCTTGGAAAGGAAGAT 1140  
Qy 1141 ATCATCATCTGCCCTCCCTTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTGCCAAG 1200  
Db 1141 ATCATCATCTGCCCTCCCTTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTGCCAAG 1200  
Qy 1201 GATCACTTAGACAAGAAAGAAAAAGCATCTGAGCCTGAAAAAGTTAAGTTCTGTCAAT 1260  
Db 1201 GATCACTTAGACAAGAAAGAAAAAGCATCTGAGCCTGAAAAAGTTAAGTTCTGTCAAT 1260  
Qy 1261 AAGTACTGTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACCATTTTGAAGAAATGG 1320  
Db 1261 AAGTACTGTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACCATTTTGAAGAAATGG 1320  
Qy 1321 TATCGTGTATTGGATTAACTGTGATACCCAACTGAAATATCATTTCCAGAAAGTTGTC 1380  
Db 1321 TATCGTGTATTGGATTAACTGTGATACCCAACTGAAATATCATTTCCAGAAAGTTGTC 1380  
Qy 1381 AAGTCTGTGATATTATTAATCAGTACAGCTCAAAATCCTTGAAAACTCCCTTTAAACTTG 1440  
Db 1381 AAGTCTGTGATATTATTAATCAGTACAGCTCAAAATCCTTGAAAACTCCCTTTAAACTTG 1440  
Qy 1441 GAAAAATGAGAAAGATGCTGGTGTTCATTTGTCAAGCTTTTCCCTCATTAATCATTTGATGAA 1500  
Db 1441 GAAAAATGAGAAAGATGCTGGTGTTCATTTGTCAAGCTTTTCCCTCATTAATCATTTGATGAA 1500  
Qy 1501 TGTCAATCACACCAACAAAGAACAGAGTATTAATTAACATCATGAGGCAATTTTGTGATGAG 1560  
Db 1501 TGTCAATCACACCAACAAAGAACAGAGTATTAATTAACATCATGAGGCAATTTTGTGATGAG 1560  
Qy 1561 AAGTTGAAAAACAATAGACTCAAGAAAAAGAAACCAACAGTATTCCTCTCAGATA 1620  
Db 1561 AAGTTGAAAAACAATAGACTCAAGAAAAAGAAACCAACAGTATTCCTCTCAGATA 1620  
Qy 1621 CTGGGACTAACAGCTTCACTGCTGTGTGTGAGGGGCCAGAGCAAGCCAAAGCTGAGAA 1680  
Db 1621 CTGGGACTAACAGCTTCACTGCTGTGTGTGAGGGGCCAGAGCAAGCCAAAGCTGAGAA 1680  
Qy 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTAATTAATAAACTGTAAAGAAAAC 1740  
Db 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTAATTAATAAACTGTAAAGAAAAC 1740  
Qy 1741 CTTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAAGTTTGGCAATGCAAGATGCA 1800  
Db 1741 CTTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAAGTTTGGCAATGCAAGATGCA 1800

Db 1741 CTTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAAGTTTGGCAATGCAAGATGCA 1800  
Qy 1801 ACCAGAGAGATCCATTTTAAAGAAACTTCTAGAAATATGACAAGATTCAAACTTAT 1860  
Db 1801 ACCAGAGAGATCCATTTTAAAGAAACTTCTAGAAATATGACAAGATTCAAACTTAT 1860  
Qy 1861 TGTCAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCTATGACAATGGGCCATTCAA 1920  
Db 1861 TGTCAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCTATGACAATGGGCCATTCAA 1920  
Qy 1921 ATGAAAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTGTGCAGAACATTG 1980  
Db 1921 ATGAAAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTGTGCAGAACATTG 1980  
Qy 1981 AGAAGTACAAATGAGGCCCTTACAAATTAAAGACAAATTCGAATGATAGATGCGTACT 2040  
Db 1981 AGAAGTACAAATGAGGCCCTTACAAATTAAAGACAAATTCGAATGATAGATGCGTACT 2040  
Qy 2041 CATCTTGAACCTTCTATATGAAGAGAAAGATTAAGAAGTTTGCAGTCATAGAAGATGAT 2100  
Db 2041 CATCTTGAACCTTCTATATGAAGAGAAAGATTAAGAAGTTTGCAGTCATAGAAGATGAT 2100  
Qy 2101 AGTGATGAGGGTGGTGATGATGAGTATTTGTGATGTTGATGATGAAGATGATTTAAAG 2160  
Db 2101 AGTGATGAGGGTGGTGATGATGAGTATTTGTGATGTTGATGATGAAGATGATTTAAAG 2160  
Qy 2161 AAACCTTTGAAACCTGATGAAACAGATGATTTCTCATGACTTTATTTTGAACAAT 2220  
Db 2161 AAACCTTTGAAACCTGATGAAACAGATGATTTCTCATGACTTTATTTTGAACAAT 2220  
Qy 2221 AAAATGTTGAAAAAGCTGGCTGAAAAACCGAATATGAAAAATGAAAACTGACCAATTA 2280  
Db 2221 AAAATGTTGAAAAAGCTGGCTGAAAAACCGAATATGAAAAATGAAAACTGACCAATTA 2280  
Qy 2281 AGAATATCCATATGGAAGCAATATATCTAGACTGAGGAATCAGACGAGGAATATCTT 2340  
Db 2281 AGAATATCCATATGGAAGCAATATATCTAGACTGAGGAATCAGACGAGGAATATCTT 2340  
Qy 2341 ACAAAAAACGACAGAGTGCATATGCGCTTTCCAGTGTGATTTACTGAAAAATGAAAAATT 2400  
Db 2341 ACAAAAAACGACAGAGTGCATATGCGCTTTCCAGTGTGATTTACTGAAAAATGAAAAATT 2400  
Qy 2401 GCTGAAGTAGAGTCAAAAGCCCAACCATTTGATTGGAGCTGACACAGCAGTGAATTCAA 2460  
Db 2401 GCTGAAGTAGAGTCAAAAGCCCAACCATTTGATTGGAGCTGACACAGCAGTGAATTCAA 2460  
Qy 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAAGTAAATTTGCCACTGAAAAATCAAT 2520  
Db 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAAGTAAATTTGCCACTGAAAAATCAAT 2520  
Qy 2521 CTGCTTATCGCTACCAAGTGGCAGAAAGGCTGATATTTAAAGAAATGTAACATTGTT 2580  
Db 2521 CTGCTTATCGCTACCAAGTGGCAGAAAGGCTGATATTTAAAGAAATGTAACATTGTT 2580  
Qy 2581 ATCCGTTATGGTCTGTCAACCAATGAAATAGCCATGGTCCAGGCCGTGTCAGCCAGA 2640  
Db 2581 ATCCGTTATGGTCTGTCAACCAATGAAATAGCCATGGTCCAGGCCGTGTCAGCCAGA 2640  
Qy 2641 GCTGATGAGAGCACTACGTCCTGTTGCTCACAGTGTTCAGAGATTATCGAACATGAG 2700  
Db 2641 GCTGATGAGAGCACTACGTCCTGTTGCTCACAGTGTTCAGAGATTATCGAACATGAG 2700  
Qy 2701 ACAGTTAATGATTTCCGAGAGAAAGATGATTAAGCTATACATTTGTTCAAAAATATG 2760  
Db 2701 ACAGTTAATGATTTCCGAGAGAAAGATGATTAAGCTATACATTTGTTCAAAAATATG 2760  
Qy 2761 AAACGAGAGAGTATGCTCATAGATTTTGGAAATTCAGATGCAAGTATATGGAAG 2820  
Db 2761 AAACGAGAGAGTATGCTCATAGATTTTGGAAATTCAGATGCAAGTATATGGAAG 2820  
Qy 2821 AAAATGAAAAACCAAGAAATATTTGCCAAGCATTAACAAGATTAACCCATCACTAATACT 2880  
Db 2821 AAAATGAAAAACCAAGAAATATTTGCCAAGCATTAACAAGATTAACCCATCACTAATACT 2880

QY 2881 TTCCTTGCACAAAGCTGAGTGTGCTAGCCTGTCTGGGGAAGATATCCATGTAAATTGAG 2940  
|||||  
Db 2881 TTCCTTGCACAAAGCTGAGTGTGCTAGCCTGTCTGGGGAAGATATCCATGTAAATTGAG 2940  
QY 2941 AAAATGCATCAGCTCAATATGACCCCAAGAACTTCAAGAACTTTACATTTGAAGAGAAAAC 3000  
|||||  
Db 2941 AAAATGCATCAGCTCAATATGACCCCAAGAACTTCAAGAACTTTACATTTGAAGAGAAAAC 3000  
QY 3001 AAAGCACTGCACAAAGAGTGTGCCGACTATCAATAATAGTGAATCATCTGCAAAATGT 3060  
|||||  
Db 3001 AAAGCACTGCACAAAGAGTGTGCCGACTATCAATAATAGTGAATCATCTGCAAAATGT 3060  
QY 3061 GGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGCTCTCAAAATA 3120  
|||||  
Db 3061 GGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGCTCTCAAAATA 3120  
QY 3121 AGGAATTTTGTAGTGTTCACAAATATTCACAAAGAAACAATACAAAAAGTGGTA 3180  
|||||  
Db 3121 AGGAATTTTGTAGTGTTCACAAATATTCACAAAGAAACAATACAAAAAGTGGTA 3180  
QY 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTTAGTATGAG 3240  
|||||  
Db 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTTAGTATGAG 3240  
QY 3241 GATTAGCACTGATTGAAGATTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300  
|||||  
Db 3241 GATTAGCACTGATTGAAGATTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300  
QY 3301 TGATTAATGATTATTATGCTACAGAACTGACATAAGAAATCAATAAATGATTGTTTA 3360  
|||||  
Db 3301 TGATTAATGATTATTATGCTACAGAACTGACATAAGAAATCAATAAATGATTGTTTA 3360  
QY 3361 CTCTG 3365  
|||||  
Db 3361 CTCTG 3365

RESULT 3  
CS113027 3380 bp DNA linear PAT 24-JUN-2005  
LOCUS CS113027 Sequence 45 from Patent WO2005054507.  
DEFINITION CS113027  
ACCESSION CS113027  
VERSION CS113027.1 GI:68224599  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1  
AUTHORS Corfe, B. and Chirakkal, H.  
TITLE Gene screen  
JOURNAL Patent: WO 2005054507-A 45 16-JUN-2005;  
University of Sheffield (GB)  
FEATURES  
source 1.3380  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 100.0%; Score 3365; DB 6; Length 3380;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCGCGGCTGAGAGCCTGTGACAACTCTCATTTGTCAAGGACAGAGCGGTAGAC 60  
|||||  
Db 1 GGGCGCGGCTGAGAGCCTGTGACAACTCTCATTTGTCAAGGACAGAGCGGTAGAC 60  
QY 61 CTTGCTTCTTAAGTGGGACGAGACAGCGGACGACATTTCACTGTCCGACAGACAA 120  
|||||  
Db 61 CTTGCTTCTTAAGTGGGACGAGACAGCGGACGACATTTCACTGTCCGACAGACAA 120

QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAATGTGGAATGGG 180  
|||||  
Db 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAATGTGGAATGGG 180  
QY 181 TATTCACAGACGAAATTTCCGCTATCTCATCTCGTCTTCAAGGCCAGGGTGAATATG 240  
|||||  
Db 181 TATTCACAGACGAAATTTCCGCTATCTCATCTCGTCTTCAAGGCCAGGGTGAATATG 240  
QY 241 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCTGCTGACAGAGGTGAAGAG 300  
|||||  
Db 241 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCTGCTGACAGAGGTGAAGAG 300  
QY 301 CAGATTCAGAGACAGTCCGCCACCTCCGGGAACATGACAGGCAATTGAATCTGCTGAGC 360  
|||||  
Db 301 CAGATTCAGAGACAGTCCGCCACCTCCGGGAACATGACAGGCAATTGAATCTGCTGAGC 360  
QY 361 ACCTTGAGAGAGGAGTCTGGCACCTTGTTGACTCGGGAATTCGTGAGGCCCTCCGG 420  
|||||  
Db 361 ACCTTGAGAGAGGAGTCTGGCACCTTGTTGACTCGGGAATTCGTGAGGCCCTCCGG 420  
QY 421 AGAACCGGACGCCCTCTGGCCGCCCTACATGAACCCCTGAGCTACAGGACTTGCCCTCT 480  
|||||  
Db 421 AGAACCGGACGCCCTCTGGCCGCCCTACATGAACCCCTGAGCTACAGGACTTGCCCTCT 480  
QY 481 CCATCGTTGAGAACGCTCATGATGAATATCTCCAATGCTGTAACCTCTTACGCCACT 540  
|||||  
Db 481 CCATCGTTGAGAACGCTCATGATGAATATCTCCAATGCTGTAACCTCTTACGCCACT 540  
QY 541 CTGCTGACAAAGCTTTCTAGTTAGAGACGCTTGGATTAAGTGCATGAGAGGAAGAACTGTTG 600  
|||||  
Db 541 CTGCTGACAAAGCTTTCTAGTTAGAGACGCTTGGATTAAGTGCATGAGAGGAAGAACTGTTG 600  
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAACCAATGAAATGAATCAGGTGTA 660  
|||||  
Db 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAACCAATGAAATGAATCAGGTGTA 660  
QY 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAACCTGTTCTGCAATTTCTGAATGTT 720  
|||||  
Db 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAACCTGTTCTGCAATTTCTGAATGTT 720  
QY 721 CTTGCTCAACACAGAAACAATGAACCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA 780  
|||||  
Db 721 CTTGCTCAACACAGAAACAATGAACCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA 780  
QY 781 AGCAATGCAGAGATTGAGAATTTATCAAAAGTTGATGCTCTCAAGTGAAGACCACTT 840  
|||||  
Db 781 AGCAATGCAGAGATTGAGAATTTATCAAAAGTTGATGCTCTCAAGTGAAGACCACTT 840  
QY 841 CTTTCAACCACAGTTCAGCCCAATCTGGAGAAAGAGTCTGGGCAATGAGATAACTCA 900  
|||||  
Db 841 CTTTCAACCACAGTTCAGCCCAATCTGGAGAAAGAGTCTGGGCAATGAGATAACTCA 900  
QY 901 TCAGATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAATCAGACACAAGTTTGCAGAA 960  
|||||  
Db 901 TCAGATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAATCAGACACAAGTTTGCAGAA 960  
QY 961 GGAAGTGTCAAGTCTTAGATGAAGTCTTGAACATTAACAGCAATGAGGCAATGATTCA 1020  
|||||  
Db 961 GGAAGTGTCAAGTCTTAGATGAAGTCTTGAACATTAACAGCAATGAGGCAATGATTCA 1020  
QY 1021 GGCAACCATGGAAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGACATCCCGGAGCCA 1080  
|||||  
Db 1021 GGCAACCATGGAAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGACATCCCGGAGCCA 1080  
QY 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCAGGCTTGAAGGGAAGAAAT 1140  
|||||  
Db 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCAGGCTTGAAGGGAAGAAAT 1140  
QY 1141 ATCATCATCTGCTCCCTACAGAGGAGTGAAGAAACAGAGTGGCTTTTACATTGCAAG 1200  
|||||  
Db 1141 ATCATCATCTGCTCCCTACAGAGGAGTGAAGAAACAGAGTGGCTTTTACATTGCAAG 1200

QY 1201 GATCACTTAGACAAGAAAAAAGCATCTGAGCCTGAAAAAGTTATAGTCTTGTCAT 1260  
|||||  
Db 1201 GATCACTTAGACAAGAAAAAAGCATCTGAGCCTGAAAAAGTTATAGTCTTGTCAT 1260  
QY 1261 AAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAAACCATTTTGAAGAAATGG 1320  
|||||  
Db 1261 AAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAAACCATTTTGAAGAAATGG 1320  
QY 1321 TATCGTGTATATGGATTAAAGTGTGATACCCCACTGAAAAATATCATTTCCAGAAAGTTGTC 1380  
|||||  
Db 1321 TATCGTGTATATGGATTAAAGTGTGATACCCCACTGAAAAATATCATTTCCAGAAAGTTGTC 1380  
QY 1381 AAGTCTGTGATATTATATTCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAAACTTG 1440  
|||||  
Db 1381 AAGTCTGTGATATTATATTCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAAACTTG 1440  
QY 1441 GAAAAATGAGAGAGATGCTGTGTTCAATTGTCAAGCTTTTCCCTCATTTATCATTTGATGAA 1500  
|||||  
Db 1441 GAAAAATGAGAGAGATGCTGTGTTCAATTGTCAAGCTTTTCCCTCATTTATCATTTGATGAA 1500  
QY 1501 TGTTCATCACACCAACAAGAAAGCAGTGTATATATACATCATGAGGCATTTTGTGATGCAG 1560  
|||||  
Db 1501 TGTTCATCACACCAACAAGAAAGCAGTGTATATATACATCATGAGGCATTTTGTGATGCAG 1560  
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACAACCCAGTATTCCTCTCTCAGATA 1620  
|||||  
Db 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACAACCCAGTATTCCTCTCTCAGATA 1620  
QY 1621 CTGGGACTAACAGCTTCACTGTGTGTTGGAGGGGCCAAGCAAGCCAAAGCTGAAGAA 1680  
|||||  
Db 1621 CTGGGACTAACAGCTTCACTGTGTGTTGGAGGGGCCAAGCAAGCCAAAGCTGAAGAA 1680  
QY 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTACTATTAAAACTGTAAAGAAAC 1740  
|||||  
Db 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTACTATTAAAACTGTAAAGAAAC 1740  
QY 1741 CTTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAAGTTGCCATTGCAGATGCA 1800  
|||||  
Db 1741 CTTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAAGTTGCCATTGCAGATGCA 1800  
QY 1801 ACCAGAGAAGATCCATTTTAAAGAGAACTTCTAGAAATATATGACAAGGATTCAACTTAT 1860  
|||||  
Db 1801 ACCAGAGAAGATCCATTTTAAAGAGAACTTCTAGAAATATATGACAAGGATTCAACTTAT 1860  
QY 1861 TGTCAAAATGAGTCCAATGTCAAGATTTTGGAACTCAACCTATGAACAATGGCCATTCAA 1920  
|||||  
Db 1861 TGTCAAAATGAGTCCAATGTCAAGATTTTGGAACTCAACCTATGAACAATGGCCATTCAA 1920  
QY 1921 ATGGAATAAAAAAGCTGCAAAAAAGAAATCGCAAAAGCAGTGTGTGCGAAACATTTG 1980  
|||||  
Db 1921 ATGGAATAAAAAAGCTGCAAAAAAGAAATCGCAAAAGCAGTGTGTGCGAAACATTTG 1980  
QY 1981 AGGAAGTACAATGAGGCCCTTCAAAATTAATGACACAAATTCGAATGATAGATGCGTATACT 2040  
|||||  
Db 1981 AGGAAGTACAATGAGGCCCTTCAAAATTAATGACACAAATTCGAATGATAGATGCGTATACT 2040  
QY 2041 CATCTTGAACCTTCTATTAATGAAGAGAAAGATGAAGTTGCAGTCAATGAAGATGAT 2100  
|||||  
Db 2041 CATCTTGAACCTTCTATTAATGAAGAGAAAGATGAAGTTGCAGTCAATGAAGATGAT 2100  
QY 2101 AGTGATGAGGGTGTGATGATGAGTATTTGATGTGATGAAGATGAGATGATTTAAAG 2160  
|||||  
Db 2101 AGTGATGAGGGTGTGATGATGAGTATTTGATGTGATGAAGATGAGATGATTTAAAG 2160  
QY 2161 AAACCTTTGAAACTGATGAACAAGATAGATTTCTCATGACTTTATTTTGAACAAT 2220  
|||||  
Db 2161 AAACCTTTGAAACTGATGAACAAGATAGATTTCTCATGACTTTATTTTGAACAAT 2220  
QY 2221 AAAATGTTGAAAAAGGCTGGCTGAAAAACCCAGAATATGAAAAATGAAAAGCTGACCAATTA 2280  
|||||  
Db 2221 AAAATGTTGAAAAAGGCTGGCTGAAAAACCCAGAATATGAAAAATGAAAAGCTGACCAATTA 2280  
QY 2281 AGAAATACCATATATGAGCAATATATCTAGACTGAGGAATCAGCAGGAATATATCTTT 2340

Db 2281 AGAAATACCATATATGAGCAATATATCTAGACTGAGGAATCAGCAGGAATAATCTTT 2340  
QY 2341 ACAAAAAACAGCACAGAGTGCATATGCGCTTTCCCAAGTGAGATTACTGAAAAATGAAAAATTT 2400  
|||||  
Db 2341 ACAAAAAACAGCACAGAGTGCATATGCGCTTTCCCAAGTGAGATTACTGAAAAATGAAAAATTT 2400  
QY 2401 GCTGAAGTAGAGTCAAAAGCCCAACATCTGATTGGAGCTGGACACAGCAGTGAATTCAAA 2460  
|||||  
Db 2401 GCTGAAGTAGAGTCAAAAGCCCAACATCTGATTGGAGCTGGACACAGCAGTGAATTCAAA 2460  
QY 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAAATTTCCGACTGAAAAATCAAT 2520  
|||||  
Db 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAAATTTCCGACTGAAAAATCAAT 2520  
QY 2521 CTGCTTATCGCTACCAAGTGGCAGAGAAAGGTCTGATATTAAGAAATGTAAATGTT 2580  
|||||  
Db 2521 CTGCTTATCGCTACCAAGTGGCAGAGAAAGGTCTGATATTAAGAAATGTAAATGTT 2580  
QY 2581 ATCCGTTATGTTCTGCTCAACCAATGAATAGCCATGTTCCAGGCCGCTGTCAGCCAGA 2640  
|||||  
Db 2581 ATCCGTTATGTTCTGCTCAACCAATGAATAGCCATGTTCCAGGCCGCTGTCAGCCAGA 2640  
QY 2641 GCTGATGAGAGCACCTACGTCCTGTTGCTCAAGTGTTCAGAGTTATCGAACATGAG 2700  
|||||  
Db 2641 GCTGATGAGAGCACCTACGTCCTGTTGCTCAAGTGTTCAGAGTTATCGAACATGAG 2700  
QY 2701 ACAGTTAATGATTTCCGAGAGAGATGATATAAAGCTATACATGTTGTTCAAAATATG 2760  
|||||  
Db 2701 ACAGTTAATGATTTCCGAGAGAGATGATATAAAGCTATACATGTTGTTCAAAATATG 2760  
QY 2761 AAACCAAGAGAGTATGTCATAGATTTTGGAAATTAACAGATGCAAAGTATATGAGAAAG 2820  
|||||  
Db 2761 AAACCAAGAGAGTATGTCATAGATTTTGGAAATTAACAGATGCAAAGTATATGAGAAAG 2820  
QY 2821 AAAATGAAAAACAAGAAATATTTGCCAAGCATTAACAAGATTAACCATCACTAATACT 2880  
|||||  
Db 2821 AAAATGAAAAACAAGAAATATTTGCCAAGCATTAACAAGATTAACCATCACTAATACT 2880  
QY 2881 TTCTCTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGAAAGATATCCATGTAATTGAG 2940  
|||||  
Db 2881 TTCTCTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGAAAGATATCCATGTAATTGAG 2940  
QY 2941 AAAATGCATCAGTCATATGACCCCAAGATTTCAAGAACTTTACATTGTAAGAGAAAAAC 3000  
|||||  
Db 2941 AAAATGCATCAGTCATATGACCCCAAGATTTCAAGAACTTTACATTGTAAGAGAAAAAC 3000  
QY 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACATCAAAATTAATGTTGTAATCATCTGCAATGT 3060  
|||||  
Db 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACATCAAAATTAATGTTGTAATCATCTGCAATGT 3060  
QY 3061 GGCCAGGCTTGGGAAACAATGATGTGACAAAGGCTTAGATTTGCTGTCTCAAAATA 3120  
|||||  
Db 3061 GGCCAGGCTTGGGAAACAATGATGTGACAAAGGCTTAGATTTGCTGTCTCAAAATA 3120  
QY 3121 AGGAATTTTGTAGTGTCTTCAAAAAATTAATCAACAAGAAACAATACAAAAAAGTGGTA 3180  
|||||  
Db 3121 AGGAATTTTGTAGTGTCTTCAAAAAATTAATCAACAAGAAACAATACAAAAAAGTGGTA 3180  
QY 3181 GAATTAACCTATCACATTTCCCAATCTTGACTAATTCAGAATGCTGTTAATTAAGTATGAG 3240  
|||||  
Db 3181 GAATTAACCTATCACATTTCCCAATCTTGACTAATTCAGAATGCTGTTAATTAAGTATGAG 3240  
QY 3241 GATTAAGCACTGATGAAGATTTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300  
|||||  
Db 3241 GATTAAGCACTGATGAAGATTTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300  
QY 3301 TGATTAATGATTCATTAATGCTACAGAACTGACATTAAGAAATCAATTAATGATGTTTA 3360  
|||||  
Db 3301 TGATTAATGATTCATTAATGCTACAGAACTGACATTAAGAAATCAATTAATGATGTTTA 3360  
QY 3361 CTCTG 3365  
|||||

```

Db          3361 CTCTG 3365

RESULT 4
LOCUS      AF095844
DEFINITION Homo sapiens melanoma differentiation associated protein-5 (MDA5)
ACCESSION AF095844
VERSION    AF095844
KEYWORDS   AF095844.1 GI:11344593
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
TITLE      1 (bases 1 to 3380)
REFERENCE  Kang,D.C., Gopalkrishnan,R.V., Wu,Q., Jankowsky,E., Pyle,A.M. and
AUTHORS    Fisher,P.B.
            mda-5: An interferon-inducible putative RNA helicase with
            double-stranded RNA-dependent ATPase activity and melanoma
            growth-suppressive properties
            Proc. Natl. Acad. Sci. U.S.A. 99 (2), 637-642 (2002)
JOURNAL    11805321
PUBMED
REFERENCE  2 (bases 1 to 3380)
AUTHORS    Kang,D.-C. and Fisher,P.B.
TITLE      Direct Submission
JOURNAL    Submitted (30-SEP-1998) Urology, Columbia University, P&S, 630
            W168th, New York, NY 10032, USA
FEATURES   Location/Qualifiers
            1..3380
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /cell_line="HO-1"
                /tissue_type="melanoma"
            1..3380
                /gene="MDA5"
                /gene="MDA5"
                /codon_start=1
                /product="melanoma differentiation associated protein-5"
                /protein_id="AAG34368.1"
                /db_xref="GI:11344594"
                /translation="MSGYSTDENFRYLISCPRAVKMYIQVEPVLVLTFLPAEYKE
                OIQTIVATSGNQAVELLSTLEKGVHLGWTFEVEALRTGSPPLARVWNPETDL
                PSPSFENAHDEYIQLINLLOPTLVKLLYDVLDKMEEBELLTIEDRNRIAAENNGN
                ESGVRELLKRIYQKENWFSAPLNLROTGNLVEQLTGSDCSSENAETENLQVDGP
                QVEEQLSTTVQPNLEKEVGMENNSSSESPADSSVSESDTSLAEGSVCLDESIGH
                NSNMGSDSGTMSDSDSEENVARASPEPELQLRPYQMEVAQPALBEGNIIICLPTGSG
                KTRVAVVYIAKDHLDKKKASEPKVIVLVNVYLVLEQLFRKEFQPLKKWYRVIGL
                DTQLKISFPEVVKSCDIIISTAQILNENSLNLENGEDAGVQSDFSLIIDECHHTNK
                EAVYNNIMRHYLMQKLKNRLKKNRPVIPLEQILGTASPGVGATQKAAEHLIK
                LCANLDAFTITVYKENLDQLKNQIQEPCCKPAIADATREDPFKEKLEIMTRIQTCCQ
                MSPMSDFGTQPEQWAIQMEKKAACKGNRKEVCAEHLRKYNELQINDTIRMIDAYT
                HLETFYNEEKDKKFAVIEDSDDEGDEYCDGEDDEDLKKPLKLDETDRFLMTLFFEE
                NNKMLKRLAENPEYENKLTKLRTIMEQYRTTESARGIIPTKTRQSAVALSQWITE
                NEKFAEVGVKAAHLIGAGHSSEFKPMTQNEQKEVISKFRTGKINLLIATVAEEGLDI
                KECNIVIRYGLVTNEIAMVQARGARADESTYLVVAHSGSVIEHETVADFREKMYKK
                AIHCVOONMKPEEYAAKLIELQOMOSIMEKKMKTKRNIAKHYNKNSPLTFLCKNSVLA
                CSGEDIHVIKRNHHVNMTPFEKELYIVRENKALQKCADYQINGEIIICKGQAWGTMM
                VHKGLDLPCLKIRNFVVFKNNSTKKQYKKWVELPITFPNLDYSECLFSDHD"

ORIGIN
Query Match      100.0%; Score 3365; DB 8; Length 3380;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db          1 GCGGCGCGGCTGAGAGCCCTGTGACACACTCGTCAATTGTACAGGACAGAGCGGTAGAC 60
            |||||||
            1 GCGGCGCGGCTGAGAGCCCTGTGACACACTCGTCAATTGTACAGGACAGAGCGGTAGAC 60
            |||||||

61 CTGCTTCTTAAGTGGGACGGACAGCGGACAGGACACGACATTTCACTGTCCCGCAGACAA 120

```

Db	61	CCTGCTTCTTAAGTGGGACGGGACAGCGGACGCAATTTCACTGTCCCGACAGACAA	120
QY	121	CAGCACCATCTGCTTGGAGAAACCCCTCTCCCTTCTCTGAGAAAGAAATGTGGAATGGG	180
Db	121	CAGCACCATCTGCTTGGAGAAACCCCTCTCCCTTCTCTGAGAAAGAAATGTGGAATGGG	180
QY	181	TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAATAATG	240
Db	181	TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAATAATG	240
QY	241	TACATCCAGGTGAGCCCTGTGCTGAGCTACCTGACCTTCTGCTGACAGAGGTGAAGAG	300
Db	241	TACATCCAGGTGAGCCCTGTGCTGAGCTACCTGACCTTCTGCTGACAGAGGTGAAGAG	300
QY	301	CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGACAGGCAAGTTGAAGTGTGCTGAGC	360
Db	301	CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGACAGGCAAGTTGAAGTGTGCTGAGC	360
QY	361	ACCTTGAGAGGAGAGTCTGGCACCTTGTGTTGACTCGGGAAATTCGTGAGAGCCCTCCGG	420
Db	361	ACCTTGAGAGGAGAGTCTGGCACCTTGTGTTGACTCGGGAAATTCGTGAGAGCCCTCCGG	420
QY	421	AGAACCCGACAGCCCTCTGGCCGCCCTACATGAACCTGAGCTCACAGGACTTGCCCTCT	480
Db	421	AGAACCCGACAGCCCTCTGGCCGCCCTACATGAACCTGAGCTCACAGGACTTGCCCTCT	480
QY	481	CCATCGTTGAGAACGCTCATGTGATGAATATCTCCAAGTGTGAACCTCCTCAGCCCACT	540
Db	481	CCATCGTTGAGAACGCTCATGTGATGAATATCTCCAAGTGTGAACCTCCTCAGCCCACT	540
QY	541	CTGTGAGACAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCATGAGAGGAACTGTTG	600
Db	541	CTGTGAGACAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCATGAGAGGAACTGTTG	600
QY	601	ACAATTGAAGACAGAAACCGGATTGCTGCTGCAAAAAACAATGAAATGAATCAGTGTGA	660
Db	601	ACAATTGAAGACAGAAACCGGATTGCTGCTGCAAAAAACAATGAAATGAATCAGTGTGA	660
QY	661	AGAGAGCTACTAAAAAGGATTTGTGCAGAAAGAAAAGTGGTTCTCTGCATTTCTGAATGTT	720
Db	661	AGAGAGCTACTAAAAAGGATTTGTGCAGAAAGAAAAGTGGTTCTCTGCATTTCTGAATGTT	720
QY	721	CTTCGTCAAAACAGGAAACAATGAACCTGTCCAAAGTTAACAGGCTCTGATTTGCTCAGAA	780
Db	721	CTTCGTCAAAACAGGAAACAATGAACCTGTCCAAAGTTAACAGGCTCTGATTTGCTCAGAA	780
QY	781	AGCAATGACAGAGATTGAGAAATTTATCACAAGTTGATGCTCAAGTGAAGAGCAACTT	840
Db	781	AGCAATGACAGAGATTGAGAAATTTATCACAAGTTGATGCTCAAGTGAAGAGCAACTT	840
QY	841	CTTTCAACCACAGTTCAAGCCAATCTGGAGAGAGGTCTGGGGCATGAGAAATTAATCTCA	900
Db	841	CTTTCAACCACAGTTCAAGCCAATCTGGAGAGAGGTCTGGGGCATGAGAAATTAATCTCA	900
QY	901	TCAGAAATCATCTTTGTCAGATTCTTCTGTAGTTTCAGAATCAGACACAAGTTTGGCAGAA	960
Db	901	TCAGAAATCATCTTTGTCAGATTCTTCTGTAGTTTCAGAATCAGACACAAGTTTGGCAGAA	960
QY	961	GGAAGTGTACAGTGTCTTAGATGAAAGTCTTGGACATAACAGCAACATGGGCACTGATTC	1020
Db	961	GGAAGTGTACAGTGTCTTAGATGAAAGTCTTGGACATAACAGCAACATGGGCACTGATTC	1020
QY	1021	GGACACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA	1080
Db	1021	GGACACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA	1080
QY	1081	GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCCTTGGAAGGGAAGAAAT	1140
Db	1081	GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCCTTGGAAGGGAAGAAAT	1140
QY	1141	ATCATCATCTGCTCCCTACAGGAGTGAAGAAAACAGAGTGGCTGTTTACATTGCCAAG	1200
Db	1141	ATCATCATCTGCTCCCTACAGGAGTGAAGAAAACAGAGTGGCTGTTTACATTGCCAAG	1200

Db 1141 ATCATCATCTGCCCTACAGGAGTGAAAAACAGAGTGGCTGTTTACATTGCCAAG 1200  
QY 1201 GATCACTTAGACAAGAAAAAGCATCTGAGCTGAAAAAGTTAAGTCTTGTCAT 1260  
Db 1201 GATCACTTAGACAAGAAAAAGCATCTGAGCTGAAAAAGTTAAGTCTTGTCAT 1260  
QY 1261 AAGGTA CTGTA GTTGA CAGCTCTTCCGCAAGAGTTCCAACCATTTTGAAGAATGG 1320  
Db 1261 AAGTA CTGTA GTTGA CAGCTCTTCCGCAAGAGTTCCAACCATTTTGAAGAATGG 1320  
QY 1321 TATCGTGTATTGGATTAAAGTGTGATACCCAACTGAAAATATCATTTCCAGAAGTTGTC 1380  
Db 1321 TATCGTGTATTGGATTAAAGTGTGATACCCAACTGAAAATATCATTTCCAGAAGTTGTC 1380  
QY 1381 AAGTCTGTGATATTATTATCAGTACAGCTCAAACTCTGAAAACTCCTTAAACTTG 1440  
Db 1381 AAGTCTGTGATATTATTATCAGTACAGCTCAAACTCTGAAAACTCCTTAAACTTG 1440  
QY 1441 GAAATGAGAGAGATGCTGGTGTTCATTTGTGAGACTTTTCCCTCATTTATCATTGAGAA 1500  
Db 1441 GAAATGAGAGAGATGCTGGTGTTCATTTGTGAGACTTTTCCCTCATTTATCATTGAGAA 1500  
QY 1501 TGTCACTCACCAACAAGAAAGCAGTGATTAATACATCATGAGGCATTAATTGATGCAG 1560  
Db 1501 TGTCACTCACCAACAAGAAAGCAGTGATTAATACATCATGAGGCATTAATTGATGCAG 1560  
QY 1561 AAGTGA AAAACAATAGACTCAAGAAAGAAAAACAACAGTGATTCCTCTCAGATA 1620  
Db 1561 AAGTGA AAAACAATAGACTCAAGAAAGAAAAACAACAGTGATTCCTCTCAGATA 1620  
QY 1621 CTGGGACTAACAGCTTCACTGTGTGGAGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
Db 1621 CTGGGACTAACAGCTTCACTGTGTGGAGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
QY 1681 CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACTGTTAAAGAAAAAC 1740  
Db 1681 CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACTGTTAAAGAAAAAC 1740  
QY 1741 CTTGATCAACTGAAAAACCAATACAGAGGCCATGCAAGAAGTTGCCATTGCAGATGCA 1800  
Db 1741 CTTGATCAACTGAAAAACCAATACAGAGGCCATGCAAGAAGTTGCCATTGCAGATGCA 1800  
QY 1801 ACCAGAGAAGATCCATTTAAAGAGAACTTTAGAAATATGACAAGATTCAAACTTAT 1860  
Db 1801 ACCAGAGAAGATCCATTTAAAGAGAACTTTAGAAATATGACAAGATTCAAACTTAT 1860  
QY 1861 TGTCAAATGAGTCCAATGTCAAGATTTTGGAACTCAACCTATGAAACAAATGGGCCATTCAA 1920  
Db 1861 TGTCAAATGAGTCCAATGTCAAGATTTTGGAACTCAACCTATGAAACAAATGGGCCATTCAA 1920  
QY 1921 ATGGA AAAAAAGCTGCAAAAAAGGAATCGCAAGAACGTGTTGTGCAGAACATTTG 1980  
Db 1921 ATGGA AAAAAAGCTGCAAAAAAGGAATCGCAAGAACGTGTTGTGCAGAACATTTG 1980  
QY 1981 AGGAAGTACAATGAGGCCCTTACAAATTAATGACAAATTCGAATGATGATGCGTATACT 2040  
Db 1981 AGGAAGTACAATGAGGCCCTTACAAATTAATGACAAATTCGAATGATGATGCGTATACT 2040  
QY 2041 CATCTTGA AACTTTCTATATAAGAGAAAGATAAGATTGCACTCATAGAAATGAT 2100  
Db 2041 CATCTTGA AACTTTCTATATAAGAGAAAGATAAGATTGCACTCATAGAAATGAT 2100  
QY 2101 AGTGATGAGGGTGTGATGATGATTTGTGATGATGATGATGATGATGATTTAAAG 2160  
Db 2101 AGTGATGAGGGTGTGATGATGATTTGTGATGATGATGATGATGATGATTTAAAG 2160  
QY 2161 AAACCTTTGAAACTGATGAAACAGATGATTTCTCATGACTTTATTTTGAACAAT 2220  
Db 2161 AAACCTTTGAAACTGATGAAACAGATGATTTCTCATGACTTTATTTTGAACAAT 2220  
QY 2221 AAAATGTTGAAAAAGCTGGCTGAAAAACCCAGAATATGAAAAAGCTGACCAAAATTA 2280  
Db 2221 AAAATGTTGAAAAAGCTGGCTGAAAAACCCAGAATATGAAAAAGCTGACCAAAATTA 2280

QY 2281 AGAATACCATAATGAGCAATATATAGACTGAGGAATCAGACGAGGAATAATCTTT 2340  
Db 2281 AGAATACCATAATGAGCAATATATAGACTGAGGAATCAGACGAGGAATAATCTTT 2340  
QY 2341 ACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGATTA CTGAAAATGAAAAATTT 2400  
Db 2341 ACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGATTA CTGAAAATGAAAAATTT 2400  
QY 2401 GCTGAAGTAGAGTCAAAAGCCCACTCTGATTGGAGCTGACACAGCAGTGAGTTCAAA 2460  
Db 2401 GCTGAAGTAGAGTCAAAAGCCCACTCTGATTGGAGCTGACACAGCAGTGAGTTCAAA 2460  
QY 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAATTTGCACTGAAAAATCAAT 2520  
Db 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAATTTGCACTGAAAAATCAAT 2520  
QY 2521 CTGCTTATCGCTACCAAGTGCGAGAGAGGCTGATATTTAAGAAATGTAACATTGTT 2580  
Db 2521 CTGCTTATCGCTACCAAGTGCGAGAGAGGCTGATATTTAAGAAATGTAACATTGTT 2580  
QY 2581 ATCCGTATGCTCTGCTCACCAATGAATAGCCATGTCAGAGCCCGCTGAGCCAGA 2640  
Db 2581 ATCCGTATGCTCTGCTCACCAATGAATAGCCATGTCAGAGCCCGCTGAGCCAGA 2640  
QY 2641 GCTGATGAGACACTACGTCCTGTTGCTCACAGTGTTCAGAGTTATCGAACATGAG 2700  
Db 2641 GCTGATGAGACACTACGTCCTGTTGCTCACAGTGTTCAGAGTTATCGAACATGAG 2700  
QY 2701 ACAGTTAATGATTTCCGAGAGAGATGATATAAGCTATACATTTGTTCAAAATATG 2760  
Db 2701 ACAGTTAATGATTTCCGAGAGAGATGATATAAGCTATACATTTGTTCAAAATATG 2760  
QY 2761 AAACGAGAGAGATGCTCATAGATTTTGGAAATTA CAGATGCAAGATATAATGGAAGAAG 2820  
Db 2761 AAACGAGAGAGATGCTCATAGATTTTGGAAATTA CAGATGCAAGATATAATGGAAGAAG 2820  
QY 2821 AAAATGAAAAACCAAGAGAAATATTGCCAAGCATTA CAAAGAAATACCATCACTAATACT 2880  
Db 2821 AAAATGAAAAACCAAGAGAAATATTGCCAAGCATTA CAAAGAAATACCATCACTAATACT 2880  
QY 2881 TTCCTTTGCAAAAACTGCAGTGTGCTAGCTGTTCTGGGGAAGTATCCATGTAATTGAG 2940  
Db 2881 TTCCTTTGCAAAAACTGCAGTGTGCTAGCTGTTCTGGGGAAGTATCCATGTAATTGAG 2940  
QY 2941 AAAATGCATCAGTCAATATGACCCCAAGATTC AAGGAACCTTTACATTTGAAGAGAAAC 3000  
Db 2941 AAAATGCATCAGTCAATATGACCCCAAGATTC AAGGAACCTTTACATTTGAAGAGAAAC 3000  
QY 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCA AATAAATGTTGTAATCATCTGCAAAATGT 3060  
Db 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCA AATAAATGTTGTAATCATCTGCAAAATGT 3060  
QY 3061 GGCCAGGCTTGGGGAACAATGATGTGCA CAAAGGCTTAGATTGCTGTCTCAAAATA 3120  
Db 3061 GGCCAGGCTTGGGGAACAATGATGTGCA CAAAGGCTTAGATTGCTGTCTCAAAATA 3120  
QY 3121 AGGAATTTTGTAGTGTGTTTCAAAAAATAATCA ACAAAGAAACAATCAAAAAAGTGGTA 3180  
Db 3121 AGGAATTTTGTAGTGTGTTTCAAAAAATAATCA ACAAAGAAACAATCAAAAAAGTGGTA 3180  
QY 3181 GAATTAACCTATCACATTTCCCAATCTTGA CTAATCAGAATGCTGTTATTTAGTGATGAG 3240  
Db 3181 GAATTAACCTATCACATTTCCCAATCTTGA CTAATCAGAATGCTGTTATTTAGTGATGAG 3240  
QY 3241 GATTAACA CTGATTGAAGATTCTTTAAATACTATCA GTTAAACATTTAATATGATTA 3300  
Db 3241 GATTAACA CTGATTGAAGATTCTTTAAATACTATCA GTTAAACATTTAATATGATTA 3300  
QY 3301 TGATTAATGTATTCATATGCTACAGAACTGACAT AAGAATCAATAAATGATTGTTTA 3360  
Db 3301 TGATTAATGTATTCATATGCTACAGAACTGACAT AAGAATCAATAAATGATTGTTTA 3360

QY 3361 CTCTG 3365  
Db 3361 CTCTG 3365

RESULT 5  
CS033863  
LOCUS CS033863 3434 bp DNA linear PAT 10-MAR-2005  
DEFINITION Sequence 3369 from Patent WO2005016962.  
ACCESSION CS033863  
VERSION CS033863.1 GI:60732714  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE  
1  
AUTHORS Abbas,A., Clark,H., Ouyang,W., Williams,M.P., Wood,W.I. and Wu,T.D.  
TITLE Compositions and methods for the treatment of immune related  
diseases  
JOURNAL Patent: WO 2005016962-A 3369 24-FEB-2005;  
Genentech, Inc. (US)  
FEATURES  
source 1.3434  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 99.9%; Score 3361.8; DB 6; Length 3434;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3363; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCGCCGGCCTGAGAGCCCTGTGACAACTTCATTTGTCAAGGCACAGAGCGGTAGAC 60  
Db 55 GCGCGCCGGCCTGAGAGCCCTGTGACAACTTCATTTGTCAAGGCACAGAGCGGTAGAC 114  
QY 61 CCGCTTCTCTAAGTGGGACGGACAGCGGACGACATTTCACCTGTGCGCAGACAA 120  
Db 115 CCGCTTCTCTAAGTGGGACGGACAGCGGACGACATTTCACCTGTGCGCAGACAA 174  
QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTGAGAAAGAGATGTCGAATGG 180  
Db 175 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTGAGAAAGAGATGTCGAATGG 234  
QY 181 TATTCCACAGACGAGATTTCCGCTATCTCATCTGCTTTCAGGGCCAGGGTGAATG 240  
Db 235 TATTCCACAGACGAGATTTCCGCTATCTCATCTGCTTTCAGGGCCAGGGTGAATG 294  
QY 241 TACATCCAGGTGAGCCCTGTGCTGACTACCTGACCTTTCTGCTGAGAGGTGAAGAG 300  
Db 295 TACATCCAGGTGAGCCCTGTGCTGACTACCTGACCTTTCTGCTGAGAGGTGAAGAG 354  
QY 301 CAGATTCAAGAGACAGTCCGCACTCCGGAAATGACAGGAGTTGAATGCTGCTGAGC 360  
Db 355 CAGATTCAAGAGACAGTCCGCACTCCGGAAATGACAGGAGTTGAATGCTGCTGAGC 414  
QY 361 ACCTTGAGAGAGGAGTGTGCACTTGTGTAAGTTCGGAATTCGTGAGGCCCTCCGG 420  
Db 415 ACCTTGAGAGAGGAGTGTGCACTTGTGTAAGTTCGGAATTCGTGAGGCCCTCCGG 474  
QY 421 AGAACCCGACAGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCAGGACTTGCCCTCT 480  
Db 475 AGAACCCGACAGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCAGGACTTGCCCTCT 534  
QY 481 CCATCGTTTGAGAAAGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAAGCCACT 540  
Db 535 CCATCGTTTGAGAAAGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAAGCCACT 594  
QY 541 CTGGTGACAAGCTTCTAGTTAGAGAGCTTGTGATAAGTGCATGAGAGGAAGTGTG 600  
Db 595 CTGGTGACAAGCTTCTAGTTAGAGAGCTTGTGATAAGTGCATGAGAGGAAGTGTG 654

QY 601 ACAATTGAGACAGAAACCGGATTGCTGCTCCAGAAAACAATGGAATGATCAGTGT 660  
Db 655 ACAATTGAGACAGAAACCGGATTGCTGCTCCAGAAAACAATGGAATGATCAGTGT 714  
QY 661 AGAGAGCTACTAAAAAGATTGTGAGAAAAGAAAAGTGTCTCTGCAATTTCTGAATGT 720  
Db 715 AGAGAGCTACTAAAAAGATTGTGAGAAAAGAAAAGTGTCTCTGCAATTTCTGAATGT 774  
QY 721 CTTGTCAAAACAGGAAACAATGAACCTTGTCCAAGAGTTAACAGAGCTGTGATGCTCAGAA 780  
Db 775 CTTGTCAAAACAGGAAACAATGAACCTTGTCCAAGAGTTAACAGAGCTGTGATGCTCAGAA 834  
QY 781 AGCAATGCAGAGATTGAGATTATACAGAGTTGATGTCCTCAAGTGAAGAGCAACTT 840  
Db 835 AGCAATGCAGAGATTGAGATTATACAGAGTTGATGTCCTCAAGTGAAGAGCAACTT 894  
QY 841 CTTTCAACCCACAGTTTCAAGCCAAATCTGAGAAAGAGGTCTGGGGCATGGAATACTCA 900  
Db 895 CTTTCAACCCACAGTTTCAAGCCAAATCTGAGAAAGAGGTCTGGGGCATGGAATACTCA 954  
QY 901 TCAGAAATCATCTTTTGACAGATTCTTGTGATTTTCAAGATCAGACACAAGTTGGCAGAA 960  
Db 955 TCAGAAATCATCTTTTGACAGATTCTTGTGATTTTCAAGATCAGACACAAGTTGGCAGAA 1014  
QY 961 GGAAGTGTCAAGCTGCTTAGATGAAGTCTTGAACATTAACAGCAACATGGGCAGTATTC 1020  
Db 1015 GGAAGTGTCAAGCTGCTTAGATGAAGTCTTGAACATTAACAGCAACATGGGCAGTATTC 1074  
QY 1021 GGCACCATGGGAAGTATTCAGATGAAGATGTGCAGCAAGAGCATCCCGGAGCCA 1080  
Db 1075 GGCACCATGGGAAGTATTCAGATGAAGATGTGCAGCAAGAGCATCCCGGAGCCA 1134  
QY 1081 GAACCTCCAGCTCAGGCTTACCMAATGGAAGTTGCCAGCCAGCCTTGGAAAGGAAGAT 1140  
Db 1135 GAACCTCCAGCTCAGGCTTACCMAATGGAAGTTGCCAGCCAGCCTTGGAAAGGAAGAT 1194  
QY 1141 ATCATCATCTGCCCTCCCTACAGGAGTGAAGAAAACAGAGTGGCTGTTTACATTGCCAAG 1200  
Db 1195 ATCATCATCTGCCCTCCCTACAGGAGTGAAGAAAACAGAGTGGCTGTTTACATTGCCAAG 1254  
QY 1201 GATCACTTAGACAAGAAAGAAAAGCATCTGAGCCTGGAAGAAAGTTAATCTTGTCAAT 1260  
Db 1255 GATCACTTAGACAAGAAAGAAAAGCATCTGAGCCTGGAAGAAAGTTAATCTTGTCAAT 1314  
QY 1261 AAGTAAGTGTAGTTGAACAGCTCTTCCGCAAGAGATTCCAAACCATTTTGAAGAATGG 1320  
Db 1315 AAGTAAGTGTAGTTGAACAGCTCTTCCGCAAGAGATTCCAAACCATTTTGAAGAATGG 1374  
QY 1321 TATCGTGTATTGATTAAAGTGTGATACCAACTGAAATATATCTTCCAGAAAGTTGTC 1380  
Db 1375 TATCGTGTATTGATTAAAGTGTGATACCAACTGAAATATATCTTCCAGAAAGTTGTC 1434  
QY 1381 AAGTCTGTGATATTATATCAGTACAGCTCAATCCTTGAAAACCTCCTTAAACTTG 1440  
Db 1435 AAGTCTGTGATATTATATCAGTACAGCTCAATCCTTGAAAACCTCCTTAAACTTG 1494  
QY 1441 GAAATGAGAGAGATGCTGCTGTTCAATTGTGCACTTTCCCTCATATTGATGAA 1500  
Db 1495 GAAATGAGAGAGATGCTGCTGTTCAATTGTGCACTTTCCCTCATATTGATGAA 1554  
QY 1501 TGTTCATCAACCAACAAGAGAGAGTGTATATAATCATCATGAGGCAATTAATTGATGAG 1560  
Db 1555 TGTTCATCAACCAACAAGAGAGAGTGTATATAATCATCATGAGGCAATTAATTGATGAG 1614  
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAACAACAGTATTTCCCTTCTCAGATA 1620  
Db 1615 AAGTTGAAAAACAATAGACTCAAGAAAGAAAACAACAGTATTTCCCTTCTCAGATA 1674  
QY 1621 CTGGGACTAACAGCTTCACTGTGTTGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
Db 1675 CTGGGACTAACAGCTTCACTGTGTTGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1734

QY 1681 CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAATAAACTGTTAAAGAAAC 1740  
| | | | |  
Db 1735 CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAATAAACTGTTAAAGAAAC 1794  
QY 1741 CTTGATCAACTGAAAAACCAATAACAGAGCCCATGCAAGAAGTTGCCATTGCAGATGCA 1800  
| | | | |  
Db 1795 CTTGATCAACTGAAAAACCAATAACAGAGCCCATGCAAGAAGTTGCCATTGCAGATGCA 1854  
QY 1801 ACCAGAGAAGATCCATTTAAAGAAAACTTCTAGAAATAATGACAAAGATTCAAACTTAT 1860  
| | | | |  
Db 1855 ACCAGAGAAGATCCATTTAAAGAAAACTTCTAGAAATAATGACAAAGATTCAAACTTAT 1914  
QY 1861 TGTCAATGAGTCCAATGTCAGATTTTGGAACTCAACCTATGAACAATGGCCATTCAA 1920  
| | | | |  
Db 1915 TGTCAATGAGTCCAATGTCAGATTTTGGAACTCAACCTATGAACAATGGCCATTCAA 1974  
QY 1921 ATGAAAAAAGCTGCAAAAAAGGAATCGCAAGAAGCGTTGTGCGAAGCATTTG 1980  
| | | | |  
Db 1975 ATGAAAAAAGCTGCAAAAAAGGAATCGCAAGAAGCGTTGTGCGAAGCATTTG 2034  
QY 1981 AGGAAGTACAATGAGGCCCTTACAATTAATGACACAATTCCGAATGATAGTCGTACT 2040  
| | | | |  
Db 2035 AGGAAGTACAATGAGGCCCTTACAATTAATGACACAATTCCGAATGATAGTCGTACT 2094  
QY 2041 CATCTGAAACTTTCTATAATGAAGAGAAAGATGAAGTTGCAGTCATAGAAGATGAT 2100  
| | | | |  
Db 2095 CATCTGAAACTTTCTATAATGAAGAGAAAGATGAAGTTGCAGTCATAGAAGATGAT 2154  
QY 2101 AGTGATGAGGGTGATGATGATGATATTGTGATGGTGAAGATGAGGATGATTTAAAG 2160  
| | | | |  
Db 2155 AGTGATGAGGGTGATGATGATGATATTGTGATGGTGAAGATGAGGATGATTTAAAG 2214  
QY 2161 AAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGAACAAT 2220  
| | | | |  
Db 2215 AAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGAACAAT 2274  
QY 2221 AAAATGTTGAAAAAGCGTGGCTGAAAAACCCAGAATATGAAAAAGCTGACCAATTA 2280  
| | | | |  
Db 2275 AAAATGTTGAAAAAGCGTGGCTGAAAAACCCAGAATATGAAAAAGCTGACCAATTA 2334  
QY 2281 AGAAATACCATTAATGAGCAATATACTAGACTGAGGAATCAGACGAGGAATAATCTTT 2340  
| | | | |  
Db 2335 AGAAATACCATTAATGAGCAATATACTAGACTGAGGAATCAGACGAGGAATAATCTTT 2394  
QY 2341 ACAAAAAACGACAGAGTGCAATATGCGCTTCCCACTGATTTACTGAAAAATGAAAAATTT 2400  
| | | | |  
Db 2395 ACAAAAAACGACAGAGTGCAATATGCGCTTCCCACTGATTTACTGAAAAATGAAAAATTT 2454  
QY 2401 GCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGAGCTGGAACACAGAGTGAGTTCAA 2460  
| | | | |  
Db 2455 GCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGAACACAGAGTGAGTTCAA 2514  
QY 2461 CCCATGACACAGAATGAACAAAAAGAGTCATTAGTAAATTTGCACTGGAATAATCAAT 2520  
| | | | |  
Db 2515 CCCATGACACAGAATGAACAAAAAGAGTCATTAGTAAATTTGCACTGGAATAATCAAT 2574  
QY 2521 CTGCTTATCGCTACCAAGTGCGACAGAAGAGGTCTGATATTAAAGAATGTAAACATTGTT 2580  
| | | | |  
Db 2575 CTGCTTATCGCTACCAAGTGCGACAGAAGAGGTCTGATATTAAAGAATGTAAACATTGTT 2634  
QY 2581 ATCCGTTATGCTCTCGTACCAATGAATATAGCCATGTCAGGCCCGTGTGAGCCAGA 2640  
| | | | |  
Db 2635 ATCCGTTATGCTCTCGTACCAATGAATATAGCCATGTCAGGCCCGTGTGAGCCAGA 2694  
QY 2641 GCTGATGAGAGCACCTACGTCCTGCTGCTCACAGTGTTCAGAGATTATCGAACATGAG 2700  
| | | | |  
Db 2695 GCTGATGAGAGCACCTACGTCCTGCTGCTCACAGTGTTCAGAGATTATCGAACATGAG 2754  
QY 2701 ACAGTTAATGATTTCCGAGAGAGATGATGTATATAAGCTATACATTGTGTCAAAATATG 2760  
| | | | |  
Db 2755 ACAGTTAATGATTTCCGAGAGAGATGATGTATATAAGCTATACATTGTGTCAAAATATG 2814  
QY 2761 AAACGAGAGAGTATGCTCATTAAGATTTTGAATTAAGATGCAAAAGTATAATGAAAAAG 2820

Db 2815 AAACGAGAGAGTATGCTCATTAAGATTTTGAAATTAACAGATGCAAAGTATAATGAAAAAG 2874  
QY 2821 AAAATGAAAAACCAAGAGAAATATTGGCCAAGCATTAACAAGATTAACCATCACTAATTA 2880  
| | | | |  
Db 2875 AAAATGAAAAACCAAGAGAAATATTGGCCAAGCATTAACAAGATTAACCATCACTAATTA 2934  
QY 2881 TTCCTTTGCAAAAACTGAGTGTGCTAGCTGTCTGCGGAAGATATCCATGTAATTTGAG 2940  
| | | | |  
Db 2935 TTCCTTTGCAAAAACTGAGTGTGCTAGCTGTCTGCGGAAGATATCCATGTAATTTGAG 2994  
QY 2941 AAAATGCATCAGTCAATATGACCCCAAGATTTCAAGAACTTTACATTGTAAGAGAAAAAC 3000  
| | | | |  
Db 2995 AAAATGCATCAGTCAATATGACCCCAAGATTTCAAGAACTTTACATTGTAAGAGAAAAAC 3054  
QY 3001 AAAGCACTGCMAAAGAGAGTGTGCCGACTATCAATAATTAATGTTGAATTCATCTGCAATGT 3060  
| | | | |  
Db 3055 AAAGCACTGCMAAAGAGAGTGTGCCGACTATCAATAATTAATGTTGAATTCATCTGCAATGT 3114  
QY 3061 GGCCAGGCTTGGGGAACAATGATGTTGTCACAAAGGCTTAGATTTGCTGCTCAAAATA 3120  
| | | | |  
Db 3115 GGCCAGGCTTGGGGAACAATGATGTTGTCACAAAGGCTTAGATTTGCTGCTCAAAATA 3174  
QY 3121 AGGAATTTTGTAGTGTGTTTCAAAAAATAATCAACAAGAAACAATACAAAAAAGTGGTA 3180  
| | | | |  
Db 3175 AGGAATTTTGTAGTGTGTTTCAAAAAATAATCAACAAGAAACAATACAAAAAAGTGGTA 3234  
QY 3181 GAATTAACCTATCACATTTCCCAATCTGACTATTCAAGATGCTGTTATTTAGTATGAG 3240  
| | | | |  
Db 3235 GAATTAACCTATCACATTTCCCAATCTGACTATTCAAGATGCTGTTATTTAGTATGAG 3294  
QY 3241 GATTAGCACTTGATTGAAGATTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300  
| | | | |  
Db 3295 GATTAGCACTTGATTGAAGATTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3354  
QY 3301 TGATTAATGTAATTCATTAATGCTACAGAACTGACATTAAGATCAATAAATGATTGTTTA 3360  
| | | | |  
Db 3355 TGATTAATGTAATTCATTAATGCTACAGAACTGACATTAAGATCAATAAATGATTGTTTA 3414  
QY 3361 CTCTG 3365  
| | | | |  
Db 3415 CTCTG 3419

RESULT 6  
CS036676 3434 bp DNA linear PAT 10-MAR-2005  
LOCUS CS036676  
DEFINITION Sequence 6182 from Patent WO2005016962.  
ACCESSION CS036676  
VERSION CS036676.1 GI:60734150  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1  
AUTHORS Abbas,A., Clark,H., Ouyang,W., Williams,M.P., Wood,W.I. and Wu,T.D.  
TITLE Compositions and methods for the treatment of immune related  
diseases  
JOURNAL Patent: WO 2005016962-A 6182 24-FEB-2005;  
Genentech, Inc. (US)  
FEATURES  
source location/Qualifiers  
1..3434  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN

Query Match 99.9%; Score 3361.8; DB 6; Length 3434;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3363; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



Db 2215 AAACCTTGAACCTGATGAAACAGATAGATTCTCATGACTTATTTTTTGAAACAAT 2274

Qy 2221 AAAATGTTGAAAAGCGTGGCTGAAAAACCCAGAAATATGAAAATGAAAAGCTGACCAATTA 2280

Db 2275 AAAATGTTGAAAAGCGTGGCTGAAAAACCCAGAAATATGAAAATGAAAAGCTGACCAATTA 2280

Qy 2281 AGAAATACCATTAATGAGCAATATATCTAGAGCTGAGGAATCAGCAGAGGAATATCTTT 2340

Db 2335 AGAAATACCATTAATGAGCAATATATCTAGAGCTGAGGAATCAGCAGAGGAATATCTTT 2394

Qy 2341 ACAAAAAACAGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400

Db 2395 ACAAAAAACAGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2454

Qy 2401 GCTGAAGTAGAGTCAAAGCCCAACCATCTGATTGGAGCTGACACAGCAGTGAAGTTCAA 2460

Db 2455 GCTGAAGTAGAGTCAAAGCCCAACCATCTGATTGAGCTGACACAGCAGTGAAGTTCAA 2514

Qy 2461 CCCATGACACAGAAATGAAACAAAAGAGTCATTAATTTCCGACTGGAAAAATCAAT 2520

Db 2515 CCCATGACACAGAAATGAAACAAAAGAGTCATTAATTTCCGACTGGAAAAATCAAT 2574

Qy 2521 CTGCTTATCGCTACACAGTGGCAGAGAAGCTGATATTAAAGAAATGTAACATTGTT 2580

Db 2575 CTGCTTATCGCTACACAGTGGCAGAGAAGCTGATATTAAAGAAATGTAACATTGTT 2634

Qy 2581 ATCCGTTATGCTCTCGTCAACCAATGAATAGCCATGCTCCAGGCCCGTGGTCCAGCCAGA 2640

Db 2635 ATCCGTTATGCTCTCGTCAACCAATGAATAGCCATGCTCCAGGCCCGTGGTCCAGCCAGA 2694

Qy 2641 GCTGATGAGAGCACCCTAGCTCCTGGTGTCTCAGAGTGGTTCAGAGTTATCGAACATGAG 2700

Db 2695 GCTGATGAGAGCACCCTAGCTCCTGGTGTCTCAGAGTGGTTCAGAGTTATCGAACATGAG 2754

Qy 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGTCTCAAATATG 2760

Db 2755 ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGTCTCAAATATG 2814

Qy 2761 AAACGAGAGGATGCTCATTAAGATTGGAATTAACAGATGCAAGATTAATGAAAAAG 2820

Db 2815 AAACGAGAGGATGCTCATTAAGATTGGAATTAACAGATGCAAGATTAATGAAAAAG 2874

Qy 2821 AAAATGAAAAACCAAGAAATATTGCAAGCATTAACAAGATTAACCACTAATAACT 2880

Db 2875 AAAATGAAAAACCAAGAAATATTGCAAGCATTAACAAGATTAACCACTAATAACT 2934

Qy 2881 TTCCCTTGCAAAAACCTGAGTGTGTAGCCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940

Db 2935 TTCCCTTGCAAAAACCTGAGTGTGTAGCCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2994

Qy 2941 AAAATGCATCAGCTCAATATGACCCCAAGATTCAGAGAACTTTACATTGTAAGAGAAAC 3000

Db 2995 AAAATGCATCAGCTCAATATGACCCCAAGATTCAGAGAACTTTACATTGTAAGAGAAAC 3054

Qy 3001 AAAGCACTGCAAAAAGAGTGTGCGCACTATCAAAATAATGTTGAATCATCTGCCAATGT 3060

Db 3055 AAAGCACTGCAAAAAGAGTGTGCGCACTATCAAAATAATGTTGAATCATCTGCCAATGT 3114

Qy 3061 GCGCAGGCTTGGGAAACAATGATGTGCAAAAAGCTTAGATTGCTGCTCTCAAAATA 3120

Db 3115 GCGCAGGCTTGGGAAACAATGATGTGCAAAAAGCTTAGATTGCTGCTCTCAAAATA 3174

Qy 3121 AGGAATTTTGTAGTGTCTTCAAAAATAATTCACAAAAGAAACAATACAAAAGTGAGTA 3180

Db 3175 AGGAATTTTGTAGTGTCTTCAAAAATAATTCACAAAAGAAACAATACAAAAGTGAGTA 3234

Qy 3181 GAATTAACCTATCACATTTCCCAATCTTGACTATTGAGAATGCTGTTAATTTAGTGATGAG 3240

Db 3235 GAATTAACCTATCACATTTCCCAATCTTGACTATTGAGAATGCTGTTAATTTAGTGATGAG 3294

Qy 3241 GATTAGCACTTGATTGAAGATTCTTTAAATATCATCAGTTAAACATTTAATATGATTA 3300

Db 3295 GATTAGCACTTGATTGAAGATTCTTTAAATATCATCAGTTAAACATTTAATATGATTA 3354

Qy 3301 TGATTAAATGATTCATTATGCTACAGAACTGACATTAAGAAATCAATAAATGATTGTTTA 3360

Db 3355 TGATTAAATGATTCATTATGCTACAGAACTGACATTAAGAAATCAATAAATGATTGTTTA 3414

Qy 3361 CTCTG 3365

Db 3415 CTCTG 3419

RESULT 7

CS042815

LOCUS

DEFINITION Sequence 3369 from Patent WO2005019258.

ACCESSION CS042815

VERSION CS042815.1 GI:61849927

KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 Abbas,A., Clark,H., Ouyang,W., Williams,P.M., Wood,W.I. and Wu,T.D. Compositions and methods for the treatment of immune related diseases

JOURNAL Patent: WO 2005019258-A 3369 03-MAR-2005;

Genentech, Inc. (US)

location/Qualifiers

source 1. 3434

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match 99.9%; Score 3361.8; DB 6; Length 3434;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3363; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCGGCGGCGCTGAGAGCCCTGTGAGACAACCTGTCATTGTGAGCAGAGCGGTAGAC 60

Db 55 GCGGCGGCGCTGAGAGCCCTGTGAGACAACCTGTCATTGTGAGCAGAGCGGTAGAC 114

Qy 61 CCTGCTTCTTAAGTGGGCGAGCGGACAGCGGACGACATTTCACCTGTCGCGAGACA 120

Db 115 CCTGCTTCTTAAGTGGGCGAGCGGACAGCGGACGACATTTCACCTGTCGCGAGACA 174

Qy 121 CAGCACCATCTGCTTGGGAGAAACCTTCTCTCTGAGAAAGAAAGATGCGAATGG 180

Db 175 CAGCACCATCTGCTTGGGAGAAACCTTCTCTCTGAGAAAGAAAGATGCGAATGG 234

Qy 181 TATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAATG 240

Db 235 TATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAATG 294

Qy 241 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCTGCTGAGAGGTGAAGAG 300

Db 295 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCTGCTGAGAGGTGAAGAG 354

Qy 301 CAGATTCAGAGACAGTCCGACCTCCGGGAACATGACAGCAGTTGAATGCTGCTGAGC 360

Db 355 CAGATTCAGAGACAGTCCGACCTCCGGGAACATGACAGCAGTTGAATGCTGCTGAGC 414

Qy 361 ACCTTGAGAGAGGAGCTTGCGACCTTGCTGTTGACTCGGGAATTCGTGAGAGCCCTCGG 420

Db 415 ACCTTGAGAGAGGAGCTTGCGACCTTGCTGTTGACTCGGGAATTCGTGAGAGCCCTCGG 474

Qy 421 AGAACCGGAGAGCCCTCTGCGCGCGCGCTACATGAACCTGAGCTACGGAATTGCCCTCT 480

Db 475 AGAACCGGAGAGCCCTCTGCGCGCGCGCTACATGAACCTGAGCTACGGAATTGCCCTCT 534

Qy 481 CCATCGTTTGAGAGCGCTCATGATGAATATCTCCAAGTGTGAACCTCTTCAGCCCACT 540

Db 535 CCATCGTTTGAGAACGCTCATGATGAATATCTCCACTGCTGAACCTCCTTCAGCCCACT 594  
QY 541 CTGCTGACACAGCTTCTAGTTAGAGACGCTCTTGGAATAAGTGCATGAGAGAACTGTG 600  
Db 595 CTGGTGACACAGCTTCTAGTTAGAGACGCTCTTGGAATAAGTGCATGAGAGAACTGTG 654  
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAACAATGAAATGAATCAGGTGTA 660  
Db 655 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAACAATGAAATGAATCAGGTGTA 714  
QY 661 AGAGAGCTACTAAAAAGATTGTGAGAAAGAAAACTGGTCTCTGCAATTTCTGAATGTT 720  
Db 715 AGAGAGCTACTAAAAAGATTGTGAGAAAGAAAACTGGTCTCTGCAATTTCTGAATGTT 774  
QY 721 CTTGCTCAAAACAGAAAACAATGAAGCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA 780  
Db 775 CTTGCTCAAAACAGAAAACAATGAAGCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA 834  
QY 781 AGCAATGACAGAGATTGAGAAATTTATCAACAAGTTGATGGTCTCAAGTGAAGAGCAACTT 840  
Db 835 AGCAATGACAGAGATTGAGAAATTTATCAACAAGTTGATGGTCTCAAGTGAAGAGCAACTT 894  
QY 841 CTTTCAACCAAGTTCAAGCCAAATCTGAGAGAGAGGTCTGGGCGATGAGAAATTAATCTCA 900  
Db 895 CTTTCAACCAAGTTCAAGCCAAATCTGAGAGAGAGGTCTGGGCGATGAGAAATTAATCTCA 954  
QY 901 TCAGAAATCATCTTTGACAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGACAGAA 960  
Db 955 TCAGAAATCATCTTTGACAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGACAGAA 1014  
QY 961 GGAAGTGTACAGCTGTAGATGAAGTCTTGACATAACAGCAACATGGGCAGTGATTCA 1020  
Db 1015 GGAAGTGTACAGCTGTAGATGAAGTCTTGACATAACAGCAACATGGGCAGTGATTCA 1074  
QY 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA 1080  
Db 1075 GGCACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA 1134  
QY 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCCTTGGAAGGGAAGAT 1140  
Db 1135 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCCTTGGAAGGGAAGAT 1194  
QY 1141 ATCATCATCTGCTCCCTACAGGAGGTGAAAAACAGAGTGGCTGTTTACATTGCCAAG 1200  
Db 1195 ATCATCATCTGCTCCCTACAGGAGGTGAAAAACAGAGTGGCTGTTTACATTGCCAAG 1254  
QY 1201 GATCACTTAGACAAAGAAAAAGCATCTGAGCCCTGAAAAAGTTAAGTTCTGTCAAT 1260  
Db 1255 GATCACTTAGACAAAGAAAAAGCATCTGAGCCCTGAAAAAGTTAAGTTCTGTCAAT 1314  
QY 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACCATTTTGAAGAAATGG 1320  
Db 1315 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACCATTTTGAAGAAATGG 1374  
QY 1321 TATCGTGTATTGATTAAGTGTGATACCAACTGAAATATATCATTTCCAGAAGTGTG 1380  
Db 1375 TATCGTGTATTGATTAAGTGTGATACCAACTGAAATATATCATTTCCAGAAGTGTG 1434  
QY 1381 AAGTCTGTGATATTATATCAGTACAGCTCAAAATCCTTGAAAACTCCCTTAAACTTG 1440  
Db 1435 AAGTCTGTGATATTATATCAGTACAGCTCAAAATCCTTGAAAACTCCCTTAAACTTG 1494  
QY 1441 GAAAAATGAGAAAGTGTGTTCAATTGTGACACTTTCCCTCATTTATCATTTGATGAA 1500  
Db 1495 GAAAAATGAGAAAGTGTGTTCAATTGTGACACTTTCCCTCATTTATCATTTGATGAA 1554  
QY 1501 TGTCAATCACACCAACAAAGAGCAGTGTATATAATCATCATGAGGCATTAATTGATGACAG 1560  
Db 1555 TGTCAATCACACCAACAAAGAGCAGTGTATATAATCATCATGAGGCATTAATTGATGACAG 1614  
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAAAGAAAACAAAACAGTGAATCCCTTCTCAGATA 1620  
Db 1615 AAGTTGAAAAACAATAGACTCAAGAAAAAGAAAACAAAACAGTGAATCCCTTCTCAGATA 1674

QY 1621 CTGGACTTAACAGCTTCACCTGTGTGTTGAGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
Db 1675 CTGGACTTAACAGCTTCACCTGTGTGTTGAGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1734  
QY 1681 CACATTTTAAACTATAGTGCCTCAATCTTGATGCAATTTACTATTAAACTGTAAAGAAAC 1740  
Db 1735 CACATTTTAAACTATAGTGCCTCAATCTTGATGCAATTTACTATTAAACTGTAAAGAAAC 1794  
QY 1741 CTTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAGTTTGCCATTGCAAGATGCA 1800  
Db 1795 CTTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAGTTTGCCATTGCAAGATGCA 1854  
QY 1801 ACCAGAGAGATTCATTTAAAGAGAACTTTAGAAATATGACAAAGATTCAAACTTAT 1860  
Db 1855 ACCAGAGAGATTCATTTAAAGAGAACTTTAGAAATATGACAAAGATTCAAACTTAT 1914  
QY 1861 TGTCAATGAGTCCAATGTCAAGTTTGAACCTCAACCTATGAACAATGGCCATTCAA 1920  
Db 1915 TGTCAATGAGTCCAATGTCAAGTTTGAACCTCAACCTATGAACAATGGCCATTCAA 1974  
QY 1921 ATGGAATAAAAGCTGCAAAAAAGAAATCGCAAGAAAGCTGTGTCAGAAACATTTG 1980  
Db 1975 ATGGAATAAAAGCTGCAAAAAAGAAATCGCAAGAAAGCTGTGTCAGAAACATTTG 2034  
QY 1981 AGGAAGTACAATGAGGCCCTCAACAATTAATGACACAATTCGAATGATAGATGCGTATACT 2040  
Db 2035 AGGAAGTACAATGAGGCCCTCAACAATTAATGACACAATTCGAATGATAGATGCGTATACT 2094  
QY 2041 CATCTTGAACCTTTCTATTAATGAAGAGAAAGATTAAGAACTTTGCACTATGAAGATGAT 2100  
Db 2095 CATCTTGAACCTTTCTATTAATGAAGAGAAAGATTAAGAACTTTGCACTATGAAGATGAT 2154  
QY 2101 AGTATGAGGGTGTGATGATGAGTATGTTGATGTTGATGAGATGAGATGATTAAAG 2160  
Db 2155 AGTATGAGGGTGTGATGATGAGTATGTTGATGTTGATGAGATGAGATGATTAAAG 2214  
QY 2161 AAACCTTTGAACCTGATGAACAGATAGATTTCTCATGACTTTATTTTGAAGAAACAAT 2220  
Db 2215 AAACCTTTGAACCTGATGAACAGATAGATTTCTCATGACTTTATTTTGAAGAAACAAT 2274  
QY 2221 AAAATGTTGAAAAAGCTGGCTGAAAAACCCAGAATATGAAAAATGAAAAATTT 2280  
Db 2275 AAAATGTTGAAAAAGCTGGCTGAAAAACCCAGAATATGAAAAATGAAAAATTT 2334  
QY 2281 AGAAATACCAATATGAGCAATATATAGACTGAGGAATCAGACGAGAAATATCTTT 2340  
Db 2335 AGAAATACCAATATGAGCAATATATAGACTGAGGAATCAGACGAGAAATATCTTT 2394  
QY 2341 ACAAAAAACAGACAGAGTGCATATGCGCTTCCAGTGAATTAAGAAAAATGAAAAATTT 2400  
Db 2395 ACAAAAAACAGACAGAGTGCATATGCGCTTCCAGTGAATTAAGAAAAATGAAAAATTT 2454  
QY 2401 GCTGAAGTAGAGTCAAGCCCAACCATGATTGAGCTGAGACAGACAGTGAAGTTCAAA 2460  
Db 2455 GCTGAAGTAGAGTCAAGCCCAACCATGATTGAGCTGAGACAGACAGTGAAGTTCAAA 2514  
QY 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAATAATTTGCACTGGAATAATCAAT 2520  
Db 2515 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAATAATTTGCACTGGAATAATCAAT 2574  
QY 2521 CTGCTTATGCTAACACAGTGGCAGAAAGAGTCTGATATTAAAGAAATGTAACATTGTT 2580  
Db 2575 CTGCTTATGCTAACACAGTGGCAGAAAGAGTCTGATATTAAAGAAATGTAACATTGTT 2634  
QY 2581 ATCCGTTATGCTCTGTCACCAATGAATAGCCATGTTCCAGGCCGTGTCAGCCAGA 2640  
Db 2635 ATCCGTTATGCTCTGTCACCAATGAATAGCCATGTTCCAGGCCGTGTCAGCCAGA 2694  
QY 2641 GCTGATGAGAGCACTAAGTCTGCTGCTTCAAGTGTTCAGAGTATTCGAACATGAG 2700  
Db 2695 GCTGATGAGAGCACTAAGTCTGCTGCTTCAAGTGTTCAGAGTATTCGAACATGAG 2754

QY 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGTAATAAGCTATACATTGTTGTTCAAAATATG 2760  
|||||  
Db 2755 ACAGTTAATGATTTCCGAGAGAAGATGATGTAATAAGCTATACATTGTTGTTCAAAATATG 2814  
QY 2761 AAACGAGAGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAGATAATGGAAG 2820  
|||||  
Db 2815 AAACGAGAGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAGATAATGGAAG 2874  
QY 2821 AAAATGAAAAACAAGAAATATTGCCAAGCATTAAGAATTAACCATCACTAATACT 2880  
|||||  
Db 2875 AAAATGAAAAACAAGAAATATTGCCAAGCATTAAGAATTAACCATCACTAATACT 2934  
QY 2881 TTCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGAAGATAATCCATGTAATTGAG 2940  
|||||  
Db 2935 TTCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGAAGATAATCCATGTAATTGAG 2994  
QY 2941 AAAATGCATCAGTCAATATGACCCCAAGAACTTTACATTGTAAGAGAAAC 3000  
|||||  
Db 2995 AAAATGCATCAGTCAATATGACCCCAAGAACTTTACATTGTAAGAGAAAC 3054  
QY 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAATAAATGTTGAATCATCTGCAAAATGT 3060  
|||||  
Db 3055 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAATAAATGTTGAATCATCTGCAAAATGT 3114  
QY 3061 GGCCAGGCTTGGGAACAATGATGTGCACAAAGGCTTAGATTTGCTGCTCAAAATA 3120  
|||||  
Db 3115 GGCCAGGCTTGGGAACAATGATGTGCACAAAGGCTTAGATTTGCTGCTCAAAATA 3174  
QY 3121 AGGAATTTTGTAGTGTTCAAAAATTAATTCACAAAGAAACAATACAAAGAGTGAGTA 3180  
|||||  
Db 3175 AGGAATTTTGTAGTGTTCAAAAATTAATTCACAAAGAAACAATACAAAGAGTGAGTA 3234  
QY 3181 GAATTAACCTATCACATTTCCCAATCTTGACTATTCAGAATGCTGTTAATTAGTATGAG 3240  
|||||  
Db 3235 GAATTAACCTATCACATTTCCCAATCTTGACTATTCAGAATGCTGTTAATTAGTATGAG 3294  
QY 3241 GATTAGCACTGATTGAAGATCTTTAAATTAATCACTAGTTAAACATTTAATGATTA 3300  
|||||  
Db 3295 GATTAGCACTGATTGAAGATCTTTAAATTAATCACTAGTTAAACATTTAATGATTA 3354  
QY 3301 TGATTAATGATTCATTAATGCTACAGAACTGACATAAAGATCAATAAATGATGTTTA 3360  
|||||  
Db 3355 TGATTAATGATTCATTAATGCTACAGAACTGACATAAAGATCAATAAATGATGTTTA 3414  
QY 3361 CTCTG 3365  
|||||  
Db 3415 CTCTG 3419

RESULT 8  
CS045628 3434 bp DNA linear PAT 22-MAR-2005  
LOCUS CS045628  
DEFINITION Sequence 6182 from Patent WO2005019258.  
ACCESSION CS045628  
VERSION CS045628.1 GI:61851939  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1  
AUTHORS Abbas,A., Clark,H., Ouyang,W., Williams,P.M., Wood,W.I. and Wu,T.D.  
TITLE Compositions and methods for the treatment of immune related  
diseases  
JOURNAL Patent: WO 2005019258-A 6182 03-MAR-2005;  
Genentech, Inc. (US)  
FEATURES  
SOURCE location/Qualifiers  
1..3434  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 99.9%; Score 3361.8; DB 6; length 3434;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3363; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GCGGCGGCGCTGAGAGCCCTGTGGAACAACCTGTCATTGTCAGGCACAGAGCGGTAGAC 60  
|||||  
Db 55 GCGGCGGCGCTGAGAGCCCTGTGGAACAACCTGTCATTGTCAGGCACAGAGCGGTAGAC 114  
QY 61 CTTGCTTCTTAAGTGGGAGCGGAGCGGACGACATTTGACCTGTCCCGACAGAA 120  
|||||  
Db 115 CTTGCTTCTTAAGTGGGAGCGGAGCGGACGACATTTGACCTGTCCCGACAGAA 174  
QY 121 CAGCACCATCTGCTTGGGAGAACCCCTCCCTTCTCTGAGAAAGAAAGATGCGAATGGG 180  
|||||  
Db 175 CAGCACCATCTGCTTGGGAGAACCCCTCCCTTCTCTGAGAAAGAAAGATGCGAATGGG 234  
QY 181 TATTCCACAGACGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAATATG 240  
|||||  
Db 235 TATTCCACAGACGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAATATG 294  
QY 241 TACATCCAGGTGAGCCTGTGCTGACTAATCTGACCTTTCCTGTCAGAGGTGAAGAG 300  
|||||  
Db 295 TACATCCAGGTGAGCCTGTGCTGACTAATCTGACCTTTCCTGTCAGAGGTGAAGAG 354  
QY 301 CAGATTGAGAGCAGTCCGCACTCCGGGAACATGACAGCAGTGAATCTGCTGAGC 360  
|||||  
Db 355 CAGATTGAGAGCAGTCCGCACTCCGGGAACATGACAGCAGTGAATCTGCTGAGC 414  
QY 361 ACCTTGAGAGAGGAGTCTGGCACTGTTGGAATCGGGAATTCGTGAGGCCCTCCGG 420  
|||||  
Db 415 ACCTTGAGAGAGGAGTCTGGCACTGTTGGAATTCGTGAGGCCCTCCGG 474  
QY 421 AGAACCGGAGCCCTCTGCGCCGCCCTACATGAACCCCTGAGCTACCGACTTGCCCTCT 480  
|||||  
Db 475 AGAACCGGAGCCCTCTGCGCCGCCCTACATGAACCCCTGAGCTACCGACTTGCCCTCT 534  
QY 481 CCATCGTTTGAGAGCGCTCATGATGAATATCTCCAAGTGTGAACCTCTCAGCCCACT 540  
|||||  
Db 535 CCATCGTTTGAGAGCGCTCATGATGAATATCTCCAAGTGTGAACCTCTCAGCCCACT 594  
QY 541 CTGCTGACAAAGCTTCTAGTTAGAGAGCTTGTGATTAAGTGAAGAGGAACTGTTG 600  
|||||  
Db 595 CTGCTGACAAAGCTTCTAGTTAGAGAGCTTGTGATTAAGTGAAGAGGAACTGTTG 654  
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGCTGACAGAAACAATGAATGAATCAGGTGTA 660  
|||||  
Db 655 ACAATTGAAGACAGAAACCGGATTGCTGCTGACAGAAACAATGAATGAATCAGGTGTA 714  
QY 661 AGAGAGCTACTAAAAAGGATTGTGACAGAAAGAAAACCTGTTCTCTGCAATTTCTGAATGTT 720  
|||||  
Db 715 AGAGAGCTACTAAAAAGGATTGTGACAGAAAGAAAACCTGTTCTCTGCAATTTCTGAATGTT 774  
QY 721 CTTGCTCAAAACAGGAAACAATGAACCTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA 780  
|||||  
Db 775 CTTGCTCAAAACAGGAAACAATGAACCTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA 834  
QY 781 AGCAATGACAGATGAGAAATTTATCAGAGTTGATGTCCTCAAGTGAAGAGCAACTT 840  
|||||  
Db 835 AGCAATGACAGATGAGAAATTTATCAGAGTTGATGTCCTCAAGTGAAGAGCAACTT 894  
QY 841 CTTTCAACACAGATTCAAGCCAATCTGAGAGAGAGGTCTGGGGCATGAGATAACTCA 900  
|||||  
Db 895 CTTTCAACACAGATTCAAGCCAATCTGAGAGAGAGGTCTGGGGCATGAGATAACTCA 954  
QY 901 TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAATCAGACACAAAGTTTGGCAGAA 960  
|||||  
Db 955 TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAATCAGACACAAAGTTTGGCAGAA 1014  
QY 961 GGAAGTGTAGCTGCTTAGATGAAAGTCTTGACATTAACAGCAACATGGGCAAGTATTCA 1020  
|||||  
Db 1015 GGAAGTGTAGCTGCTTAGATGAAAGTCTTGACATTAACAGCAACATGGGCAAGTATTCA 1074

QY 1021 GGCACCATGGGAAGTATTTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCCGAGCCA 1080  
|||||  
Db 1075 GGCACCATGGGAAGTATTTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCCGAGCCA 1134  
QY 1081 GAACTCCAGCTCAGGCGCTTACCAAATGGAAGTTGCCCAAGCCAGCTTGGAAGGGAAGAT 1140  
|||||  
Db 1135 GAACTCCAGCTCAGGCGCTTACCAAATGGAAGTTGCCCAAGCCAGCTTGGAAGGGAAGAT 1194  
QY 1141 ATCATCATCTGCCCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTATTACATTGGCCAAG 1200  
|||||  
Db 1195 ATCATCATCTGCCCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTATTACATTGGCCAAG 1254  
QY 1201 GATCACTTAGACAAGAAGAAAAAGCATCTGAGCTGGAAAAAGTTAAGTTCTTGTCAT 1260  
|||||  
Db 1255 GATCACTTAGACAAGAAGAAAAAGCATCTGAGCTGGAAAAAGTTAAGTTCTTGTCAT 1314  
QY 1261 AAGTACTGTAGTGAACAGCTCTTCCGCAAGAGTTCCAACCATTTTGAAGAATGG 1320  
|||||  
Db 1315 AAGTACTGTAGTGAACAGCTCTTCCGCAAGAGTTCCAACCATTTTGAAGAATGG 1374  
QY 1321 TATCGTGTATTGGATTAAAGTGTGATACCCAACTGAAAAATATCATTTCCAGAAGTTGTC 1380  
|||||  
Db 1375 TATCGTGTATTGGATTAAAGTGTGATACCCAACTGAAAAATATCATTTCCAGAAGTTGTC 1434  
QY 1381 AAGTCTGTGATATTATATCATGACAGCTCAATCCTTGAAAACTCCCTTTAAACTTG 1440  
|||||  
Db 1435 AAGTCTGTGATATTATATCATGACAGCTCAATCCTTGAAAACTCCCTTTAAACTTG 1494  
QY 1441 GAAAAATGAGAAGATGCTGCTGTTCAATTGTCAAGCTTTCCCTCATATTCATTGATGAA 1500  
|||||  
Db 1495 GAAAAATGAGAAGATGCTGCTGTTCAATTGTCAAGCTTTCCCTCATATTCATTGATGAA 1554  
QY 1501 TGTTCATCACACCAACAAGAGCAGTGTATATTAACATCATGAGGCATATTGATGCAG 1560  
|||||  
Db 1555 TGTTCATCACACCAACAAGAGCAGTGTATATTAACATCATGAGGCATATTGATGCAG 1614  
QY 1561 AAGTGAAGAAAAATAGACTCAAGAAAGAAAAACCAAGTATTCCTTCTCAGATA 1620  
|||||  
Db 1615 AAGTGAAGAAAAATAGACTCAAGAAAGAAAAACCAAGTATTCCTTCTCAGATA 1674  
QY 1621 CTGGGACTAACAGCTTCACTGCTGTTGGAGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
|||||  
Db 1675 CTGGGACTAACAGCTTCACTGCTGTTGGAGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1734  
QY 1681 CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACTGTTTAAAGAAAC 1740  
|||||  
Db 1735 CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACTGTTTAAAGAAAC 1794  
QY 1741 CTTGATCAACTGAAAAACCAATACAGAGGCCATGCAAGAAGTTGCCATTGCAAGATGCA 1800  
|||||  
Db 1795 CTTGATCAACTGAAAAACCAATACAGAGGCCATGCAAGAAGTTGCCATTGCAAGATGCA 1854  
QY 1801 ACCAGAGAAGATCCATTTTAAAGAGAACTTCTAGAAATTAATGACAAGAGTTCAAACTTAT 1860  
|||||  
Db 1855 ACCAGAGAAGATCCATTTTAAAGAGAACTTCTAGAAATTAATGACAAGAGTTCAAACTTAT 1914  
QY 1861 TGTCAAATGAGTCCCAATGTCAGATTTTGGAACTCAACCCATATGAACAAATGGGCCATTCAA 1920  
|||||  
Db 1915 TGTCAAATGAGTCCCAATGTCAGATTTTGGAACTCAACCCATATGAACAAATGGGCCATTCAA 1974  
QY 1921 ATGGAAGAAAAAGCTGCAAAAAAGAAATGCGAAAGAAAGCTGTTGTGCAAGAACATTGG 1980  
|||||  
Db 1975 ATGGAAGAAAAAGCTGCAAAAAAGAAATGCGAAAGAAAGCTGTTGTGCAAGAACATTGG 2034  
QY 1981 AGGAAGTACAATGAGGCCCTTACAATTAATGACACAATTCGAATGATGATGCGTATACT 2040  
|||||  
Db 2035 AGGAAGTACAATGAGGCCCTTACAATTAATGACACAATTCGAATGATGATGCGTATACT 2094  
QY 2041 CATCTTGAAGCTTTCTATAATGAAGAGAAAGATAGAAGTTTGACGTGATAGAAGATGAT 2100  
|||||  
Db 2095 CATCTTGAAGCTTTCTATAATGAAGAGAAAGATAGAAGTTTGACGTGATAGAAGATGAT 2154  
QY 2101 AGTGATGAGGCTGTGATGATGAGTATTGTGATGCTGATGAAGATGAGGATGATTAAAG 2160

Db 2155 AGTGATGAGGCTGTGATGATGAGTATTGTGATGATGAAGATGAGGATGATTAAAG 2214  
QY 2161 AAACCTTTGAAACTGATGAAGACAGATAGATTCTCATGACTTTATTTTGAAGAAACAAT 2220  
|||||  
Db 2215 AAACCTTTGAAACTGATGAAGACAGATAGATTCTCATGACTTTATTTTGAAGAAACAAT 2274  
QY 2221 AAAATGTTGAAAAAGCTGGCTGAAAAACCGAATATGAAAAATGAAAAAGCTGACCAATTA 2280  
|||||  
Db 2275 AAAATGTTGAAAAAGCTGGCTGAAAAACCGAATATGAAAAATGAAAAAGCTGACCAATTA 2334  
QY 2281 AGAATATCCATTAATGAGCAATATATCTAGACTGAGGAATCAGCAGGAATAATCTTT 2340  
|||||  
Db 2335 AGAATATCCATTAATGAGCAATATATCTAGACTGAGGAATCAGCAGGAATAATCTTT 2394  
QY 2341 ACAAAAAACAGACAGAGTGCATATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATT 2400  
|||||  
Db 2395 ACAAAAAACAGACAGAGTGCATATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATT 2454  
QY 2401 GCTGAAGTAGAGTCAAGCCCAACCATCTGATTGGAGCTGGAACACAGCAGTGAAGTTCAAA 2460  
|||||  
Db 2455 GCTGAAGTAGAGTCAAGCCCAACCATCTGATTGGAGCTGGAACACAGCAGTGAAGTTCAAA 2514  
QY 2461 CCCATGACACAGATGAACAAAAAGAGTCAATTAAGTAATTTGCACTGGAATAATCAAT 2520  
|||||  
Db 2515 CCCATGACACAGATGAACAAAAAGAGTCAATTAAGTAATTTGCACTGGAATAATCAAT 2574  
QY 2521 CTGCTTATCGCTACACAGTGGCAGAGAGGTCTGATATTAAAGATGTAACATTGTT 2580  
|||||  
Db 2575 CTGCTTATCGCTACACAGTGGCAGAGAGGTCTGATATTAAAGATGTAACATTGTT 2634  
QY 2581 ATCCGTTATGCTCTGCTCAACCAATGAATAGCCATGCTCAGGCCCGTGTGAGCCAGA 2640  
|||||  
Db 2635 ATCCGTTATGCTCTGCTCAACCAATGAATAGCCATGCTCAGGCCCGTGTGAGCCAGA 2694  
QY 2641 GCTGATGAGAGCACCTACGCTGCTGCTTCAAGTGTTCAGAGTTATCGAACATGAG 2700  
|||||  
Db 2695 GCTGATGAGAGCACCTACGCTGCTGCTTCAAGTGTTCAGAGTTATCGAACATGAG 2754  
QY 2701 ACAGTTAATGATTTCCGAGAGAGATGATGTAATAAGCTATACATTGTCTCAAAATATG 2760  
|||||  
Db 2755 ACAGTTAATGATTTCCGAGAGAGATGATGTAATAAGCTATACATTGTCTCAAAATATG 2814  
QY 2761 AAACGAGAGAGTATGCTCATAGATTTTGAATTACAGATGCAAGATAATGGAAG 2820  
|||||  
Db 2815 AAACGAGAGAGTATGCTCATAGATTTTGAATTACAGATGCAAGATAATGGAAG 2874  
QY 2821 AAAATGAAGAACCAAGAGAAATATTGCCAAGCATTAACAAGATTAACCATCACTAATACT 2880  
|||||  
Db 2875 AAAATGAAGAACCAAGAGAAATATTGCCAAGCATTAACAAGATTAACCATCACTAATACT 2934  
QY 2881 TTCCTTTGCAAAAACCTGCAAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940  
|||||  
Db 2935 TTCCTTTGCAAAAACCTGCAAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2994  
QY 2941 AAAATGCATCAGCTCAATATGACCCCAAGATTTCAAGGAACCTTACATTGTAAGAGAAAC 3000  
|||||  
Db 2995 AAAATGCATCAGCTCAATATGACCCCAAGATTTCAAGGAACCTTACATTGTAAGAGAAAC 3054  
QY 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGTTGAATCATCTGCAAAATGT 3060  
|||||  
Db 3055 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGTTGAATCATCTGCAAAATGT 3114  
QY 3061 GGCCAGGCTTGGGGAACAATGATGTTGCAAAAAGGCTTAGATTGGCTGTCTCAAAATA 3120  
|||||  
Db 3115 GGCCAGGCTTGGGGAACAATGATGTTGCAAAAAGGCTTAGATTGGCTGTCTCAAAATA 3174  
QY 3121 AGGAATTTGTAGTGTGTTTCAAAAATTAATTCACAAAGAAACAATACAAAAGTGGGTA 3180  
|||||  
Db 3175 AGGAATTTGTAGTGTGTTTCAAAAATTAATTCACAAAGAAACAATACAAAAGTGGGTA 3234  
QY 3181 GAAATTAACCTATCACATTTCCCAATCTTGACTATTCAGAATGCTGTTATTAGTGATGAG 3240  
|||||

Db 3235 GAATTACCTATCACATTTCCCAATCTTGACTATTGAGAATGCTGTTATTAGTGATGAG 3294

Qy 3241 GATTAGCACTTGATTGAAGATTCTTTAAATACATTCAGTTAAACATTTAATATGATTA 3300

Db 3295 GATTAGCACTTGATTGAAGATTCTTTAAATACATTCAGTTAAACATTTAATATGATTA 3354

Qy 3301 TGATTAAATGATTATTCATTATGCTACAGAACTGACATTAAGAAATCAATAAAATGATTGTTTA 3360

Db 3355 TGATTAAATGATTATTCATTATGCTACAGAACTGACATTAAGAAATCAATAAAATGATTGTTTA 3414

Qy 3361 CTCTG 3365

Db 3415 CTCTG 3419

RESULT 9

AX300832

LOCUS AX300832 3372 bp DNA linear PAT 16-APR-2005

DEFINITION Sequence 1 from Patent WO0185955.

AX300832

VERSION AX300832.1 GI:17382110

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Bahr, G., Cocude, C. and Capron, A.

AUTHORS Rh16 polypeptides and its fragments and polynucleotides encoding

TITLE said polypeptides and therapeutic uses

JOURNAL Patent: WO 0185955-A 1 15-NOV-2001;

FEATURES

source location/Qualifiers

1

1.3372

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

155..3232

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD13049.1"

/db\_xref="GI:17382111"

/translation="MSGYSTDENFRYLISCFRAVVKMYIQVEPVLDTLFLPAEYKE

QIQTAVTSGNMQAVELLSTLEKGVHLEWTFREFEALRTGSPLAARYMNPETDL

PSPFENAHDEYLQNLNLQPTLVDKLVLDKCMEEELLTIEDRNRIAAEENGN

ESGRELKRIIVQKENVFSAFLNLRQTNMELVQELTSGDSNSVAEIEINLSOVDGP

QVEEQLSTTVQPNLEKEVWGMENSSSESPADSSVSESSTSLAEGSVSLDESLGH

NSNMGSDSGTMSDSEENVAARASPEPELQLRPYQMEVAQPALEGNIIICLPTGSG

KTRVAVYIAKDHLDDKKKASEPGKIVLVNVLVLEQLFRKEFQPLKMYRVLGLSG

DTQKISPEPVVKSCDIILISTAQILENSLNLNENGEDAGVQLSDSPFIIDECHTNK

EAVYNNIMRHYLMQKLKNRLKKNRPVLPQILGLTASPGVGATKQAKAEHILK

LCANLDAFTIKTVKENDLQLNQIQEPCKRAIADATREDPFKEKLEIMTRIQTCCQ

MSPMSDFGTOPYEOWAIQMEKKAKEGNRKESVCAEHLRKYNKALQINDTIRMIDAYT

HLETYPNEEKDKKRAVIDEDSDEGDEYCDGDEDEDDLKPLKLDETDRFLMTLFFE

NNKMLKRLAENPEYENELTKLRNTIMEQYRTESARGIIFTKROSAYALQOWITE

NEKFAEVGVKAHHLIGAGHSSEFKPMTONEQKEVISKFRGKINILTIATVAEEGLDI

KECNIVIRYGLVTNETAMVQAGRARADESTYVLVAHSGSGVIERETVNDFREKMMYK

AHCVQNMKPEEYAHKILELOMQSIMEKKMTKRNIAKHYNKPSLITFLCKNCSVLA

CSGEDIHVIEKMHVNMTPPEFKELYIVRENKALQKCADYQINGEITICKGQAMGTMM

VHKGIDLPLKIRNPFVVVFKNSTKKQYKMWELPITFPNLDYSECCLPFSDDED"

ORIGIN

Query Match 99.2%; Score 3338.8; DB 6; Length 3372;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3343; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 15 GAGCCCTGTGACAACCTGTCATTGTTCAGGCACAGAGCGGTAGACCCTGCTCTCTAAG 74

Db 1 GGGCCCTGTGACAACCTGTCATTGTTCAGGCACAGAGCGGTAGACCCTGCTCTCTAAG 60

Qy 75 TGGGACGGGACAGCGGACGACATTTACCTGTCCCGACAGACAAAGACCATCTGCT 134

Db 61 TGGGACGGGACAGCGGACGACATTTACCTGTCCCGACAGACAAAGACCATCTGCT 120

Qy 135 TGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGGTATTCCACAGACGA 194

Db 121 TGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGGTATTCCACAGACGA 180

Qy 195 GAATTTCCGCTATCTCATCTCGTCTTCAAGGGCCAGGGTGAAAATGTACATCCAGGTGA 254

Db 181 GAATTTCCGCTATCTCATCTCGTCTTCAAGGGCCAGGGTGAAAATGTACATCCAGGTGA 240

Qy 255 GCCTGTGCTGACTACCTGACCTTTCTGCTGACAGAGGTGAAGGACAGATTCAAGGAC 314

Db 241 GCCTGTGCTGACTACCTGACCTTTCTGCTGACAGAGGTGAAGGACAGATTCAAGGAC 300

Qy 315 AGTCGCCACCTCCGGGACATGACAGGAGTTGAATGCTGCTGAGCACCTTGAGAGAGG 374

Db 301 AGTCGCCACCTCCGGGACATGACAGGAGTTGAATGCTGCTGAGCACCTTGAGAGAGG 360

Qy 375 AGTCGGCACCTTGGTTGACTCCGGGAATTCGTGAGGCCCTCCGAGAACCGGACGCC 434

Db 361 AGTCGGCACCTTGGTTGACTCCGGGAATTCGTGAGGCCCTCCGAGAACCGGACGCC 420

Qy 435 TCTGCGCCGCCGCTACATGAACCCCTGAGCTCACGGAATTGCCCCCTCCATCGTTGAGAA 494

Db 421 TCTGCGCCGCCGCTACATGAACCCCTGAGCTCACGGAATTGCCCCCTCCATCGTTGAGAA 480

Qy 495 CGCTCATGATGAATATCTTCCAACCTGCTGAACCTCTTACGCCCACTCTGGTGACAAAGCT 554

Db 481 CGCTCATGATGAATATCTTCCAACCTGCTGAACCTCTTACGCCCACTCTGGTGACAAAGCT 540

Qy 555 TCTAGTTAGAGACGCTTGGATTAAGTCATGAGAGGAACTGTTGACAAATTGAAGACAG 614

Db 541 TCTAGTTAGAGACGCTTGGATTAAGTCATGAGAGGAACTGTTGACAAATTGAAGACAG 600

Qy 615 AAACCGGATTGCTGCTGACAGAAACAATGAAATGAATCAGGTGTAAGAGACTACTAA 674

Db 601 AAACCGGATTGCTGCTGACAGAAACAATGAAATGAATCAGGTGTAAGAGACTACTAA 660

Qy 675 AAGATTGTGCAGAAAGAAAACTGGTCTCTGCAATTTCTGAATGTTCTTGTCAAAACAG 734

Db 661 AAGATTGTGCAGAAAGAAAACTGGTCTCTGCAATTTCTGAATGTTCTTGTCAAAACAG 720

Qy 735 AAACATGAACCTGTCCAGAGTTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGAT 794

Db 721 AAACATGAACCTGTCCAGAGTTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGAT 780

Qy 795 TGAGATTATCACAAAGTTGATGCTCCTCAAGTGAAGAGCAACTTCTTCAACACAGT 854

Db 781 TGAGATTATCACAAAGTTGATGCTCCTCAAGTGAAGAGCAACTTCTTCAACACAGT 840

Qy 855 TCAGCAAATCTGAGAGAGAGTCTGGGACATGAGAAATTAATCATCAGAAATCATCTTT 914

Db 841 TCAGCAAATCTGAGAGAGAGTCTGGGACATGAGAAATTAATCATCAGAAATCATCTTT 900

Qy 915 TGCAGATTCTTCTGATTTCAGATCAGACACAAAGTTTGGCAGAAAGAGTTCAGCTG 974

Db 901 TGCAGATTCTTCTGATTTCAGATCAGACACAAAGTTTGGCAGAAAGAGTTCAGCTG 960

Qy 975 CTTAGATGAAGAAGTCTTGACATAACAGCAACATGGGCAAGTATTCAGGCAACCATGGGAAG 1034

Db 961 CTTAGATGAAGAAGTCTTGACATAACAGCAACATGGGCAAGTATTCAGGCAACCATGGGAAG 1020

Qy 1035 TGATTAGATGAAGAGATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAG 1094

Db 1021 TGATTAGATGAAGAGATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAG 1080

Qy 1095 GCCTTACCAAATGAAGTGGCCAGCCAGCTTGGAAAGGAAAGATATCATCTGCGCT 1154

Db 1081 GCCTTACCAAATGAAGTGGCCAGCCAGCTTGGAAAGGAAAGATATCATCTGCGCT 1140

Qy 1155 CCCTACAGGAGTGAAGAAACAGAGTGGCTGTTACATTGCCAAGGATCACTTAGACAA 1214

Db 1141 CCCTACAGGAGTGAAGAAACAGAGTGGCTGTTACATTGCCAAGGATCACTTAGACAA 1200

QY 1215 GAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTGTCAATAAGTACTGCTAGT 1274  
| | | | |  
Db 1201 GAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTGTCAATAAGTACTGCTAGT 1260  
QY 1275 TGAACAGCTCTTCCGCAAGAGAGTTCCAACCAATTTTGAAGAAATGATCGTGTATTGG 1334  
| | | | |  
Db 1261 TGAACAGCTCTTCCGCAAGAGAGTTCCAACCAATTTTGAAGAAATGATCGTGTATTGG 1320  
QY 1335 ATTAAGTGTGATATACCCAACTGAAAAATATCATTTTCCAGAAGTTGTCAAGTCCGTGATAT 1394  
| | | | |  
Db 1321 ATTAAGTGTGATATACCCAACTGAAAAATATCATTTTCCAGAAGTTGTCAAGTCCGTGATAT 1380  
QY 1395 TATTATCAGTACAGCTCAAACTCTTGAAAACTCCCTCTTAACTTGAAAAATGAGAGA 1454  
| | | | |  
Db 1381 TATTATCAGTACAGCTCAAACTCTTGAAAACTCCCTCTTAACTTGAAAAATGAGAGA 1440  
QY 1455 TGCTGTGTTCAATTTGTACAGCTTTTCCCTCATTTATCATGTGATGAATGTCAACCAA 1514  
| | | | |  
Db 1441 TGCTGTGTTCAATTTGTACAGCTTTTCCCTCATTTATCATGTGATGAATGTCAACCAA 1500  
QY 1515 CAAGAAGCAGTGTATATAATACATCATGAGCATTATTGATGCAGAAGTTGAAAACAA 1574  
| | | | |  
Db 1501 CAAGAAGCAGTGTATATAATACATCATGAGCATTATTGATGCAGAAGTTGAAAACAA 1560  
QY 1575 TAGACTCAAGAAAAAGAAACAACCAAGTATCCCTTCTCAAGTACTGGACTTAACAGC 1634  
| | | | |  
Db 1561 TAGACTCAAGAAAAAGAAACAACCAAGTATCCCTTCTCAAGTACTGGACTTAACAGC 1620  
QY 1635 TTCACTGTGTGTGGAGGGGCCAAGAGCAAGCCAAAGCTGAAGAACACATTTTAAACT 1694  
| | | | |  
Db 1621 TTCACTGTGTGTGGAGGGGCCAAGAGCAAGCCAAAGCTGAAGAACACATTTTAAACT 1680  
QY 1695 ATGTGCCAATCTTGATGCATTTAATAAACTGTTAAAGAAAACCTTGATCAACTGAA 1754  
| | | | |  
Db 1681 ATGTGCCAATCTTGATGCATTTAATAAACTGTTAAAGAAAACCTTGATCAACTGAA 1740  
QY 1755 AAACCAAAATACAGAGCCATGCAAGAAAGTTGCCATTCAGATGCAACCAAGAGAAGATCC 1814  
| | | | |  
Db 1741 AAACCAAAATACAGAGCCATGCAAGAAAGTTGCCATTCAGATGCAACCAAGAGAAGATCC 1800  
QY 1815 ATTTAAGAGAACTTCTAGAAATATGACACAGGATTCAAACTTATTGTCAATGAGTCC 1874  
| | | | |  
Db 1801 ATTTAAGAGAACTTCTAGAAATATGACACAGGATTCAAACTTATTGTCAATGAGTCC 1860  
QY 1875 AATGTCAATTTTGGAACTCAACCTTATGAACAATGGCCATTCAATGAAAAAAAGC 1934  
| | | | |  
Db 1861 AATGTCAATTTTGGAACTCAACCTTATGAACAATGGCCATTCAATGAAAAAAAGC 1920  
QY 1935 TGCAAAAAAGAAATCGCAAGAAAGTGTGTGTGCAAGAACATTTGAGGAAGTACAATGA 1994  
| | | | |  
Db 1921 TGCAAAAAAGAAATCGCAAGAAAGTGTGTGTGCAAGAACATTTGAGGAAGTACAATGA 1980  
QY 1995 GGCCCTACAATTAATGACACAAATTCGAATGATAGATGCGTACTCATCTTGAACCTT 2054  
| | | | |  
Db 1981 GGCCCTACAATTAATGACACAAATTCGAATGATAGATGCGTACTCATCTTGAACCTT 2040  
QY 2055 CTATATGAAGAGAAAGATTAAGAGTTTGCAGTCAATGAAGATGATAGTGAAGGTTGG 2114  
| | | | |  
Db 2041 CTATATGAAGAGAAAGATTAAGAGTTTGCAGTCAATGAAGATGATAGTGAAGGTTGG 2100  
QY 2115 TGATGATGAGTATTGTGATGGTGAAGATGAGATGATTTAAAGAAACCTTTGAAACT 2174  
| | | | |  
Db 2101 TGATGATGAGTATTGTGATGGTGAAGATGAGATGATTTAAAGAAACCTTTGAAACT 2160  
QY 2175 GGATGAAACAGATAGATTTTCTCATGACTTTATTTTGAACAAATAAATGTTGAAAAAG 2234  
| | | | |  
Db 2161 GGATGAAACAGATAGATTTTCTCATGACTTTATTTTGAACAAATAAATGTTGAAAAAG 2220  
QY 2235 GCTGCTGAACCAAGCAATATGAACAAAGCTGACCAATTAAGAAATACATAAT 2294  
| | | | |  
Db 2221 GCTGCTGAACCAAGCAATATGAACAAAGCTGACCAATTAAGAAATACATAAT 2280

QY 2295 GGAGCAATATACTAGACTGAGGAATCAGACAGGAATAATCTTTACAAAAACAGACA 2354  
| | | | |  
Db 2281 GGAGCAATATACTAGACTGAGGAATCAGACAGGAATAATCTTTACAAAAACAGACA 2340  
QY 2355 GAGTCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGAGT 2414  
| | | | |  
Db 2341 GAGTCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGAGT 2400  
QY 2415 CAAAGCCCAACCATCTGATTTGAGCTGGACACAGAGTGAAGTTCAAAACCAATGACACAGA 2474  
| | | | |  
Db 2401 CAAAGCCCAACCATCTGATTTGAGCTGGACACAGAGTGAAGTTCAAAACCAATGACACAGA 2460  
QY 2475 TGAACAAAAAGAGTCAATTAAGTAAATTTCCACTGAAAAAATCAATCTGCTTATCGCTAC 2534  
| | | | |  
Db 2461 TGAACAAAAAGAGTCAATTAAGTAAATTTCCACTGAAAAAATCAATCTGCTTATCGCTAC 2520  
QY 2535 CACAGTGGCAGAGAAGTCTGATATTAAAGATGTAAACATTTGTTATCCGTTATGCTCT 2594  
| | | | |  
Db 2521 CACAGTGGCAGAGAAGTCTGATATTAAAGATGTAAACATTTGTTATCCGTTATGCTCT 2580  
QY 2595 CGTCAACCAATGAAATAGCATGTGTCAGGCCCCGTGTCAGGCCAGAGCTGATGACAGCAC 2654  
| | | | |  
Db 2581 CGTCAACCAATGAAATAGCATGTGTCAGGCCCCGTGTCAGGCCAGAGCTGATGACAGCAC 2640  
QY 2655 CTACGTCTGTGTTCTCAGTGTGTTTCAAGAGTTATCGAACATGAGACAGTTAATGATTT 2714  
| | | | |  
Db 2641 CTACGTCTGTGTTCTCAGTGTGTTTCAAGAGTTATCGAACATGAGACAGTTAATGATTT 2700  
QY 2715 CCGAGAGAAGATGATGTATAAGCTATACATTGTCTCAAAATATGAACCAAGAGAGTA 2774  
| | | | |  
Db 2701 CCGAGAGAAGATGATGTATAAGCTATACATTGTCTCAAAATATGAACCAAGAGAGTA 2760  
QY 2775 TGCTCATAGATTTTGGAAATTAACAGATGCAAGTATAATGAAAAAGAAAAATGAAAAACCA 2834  
| | | | |  
Db 2761 TGCTCATAGATTTTGGAAATTAACAGATGCAAGTATAATGAAAAAGAAAAATGAAAAACCA 2820  
QY 2835 GAGAAATATTGCCAAGCATTAACAAGATAACCACTCACTAATACTTTCTTGCAAAAA 2894  
| | | | |  
Db 2821 GAGAAATATTGCCAAGCATTAACAAGATAACCACTCACTAATACTTTCTTGCAAAAA 2880  
QY 2895 CTGCAGTGTGCTAGCCTGTCTGGGGAAGATATCCATGTAATTGAGAAAAATGCATCACGT 2954  
| | | | |  
Db 2881 CTGCAGTGTGCTAGCCTGTCTGGGGAAGATATCCATGTAATTGAGAAAAATGCATCACGT 2940  
QY 2955 CAATATGACCCCAAGATTCAAGAACTTTACATTGTAAAGAAAAACAACCACTGCAAAA 3014  
| | | | |  
Db 2941 CAATATGACCCCAAGATTCAAGAACTTTACATTGTAAAGAAAAACAACCACTGCAAAA 3000  
QY 3015 GAAAGTGCCTGACTATCAATAAATGTTGAATCATCTGCAATGTGGCCAGGCTTGGGG 3074  
| | | | |  
Db 3001 GAAAGTGCCTGACTATCAATAAATGTTGAATCATCTGCAATGTGGCCAGGCTTGGGG 3060  
QY 3075 AACATGATGTGTCACAAAGGCTTAGATTGCTGCTCAAAATTAAGGAATTTTGTAGT 3134  
| | | | |  
Db 3061 AACATGATGTGTCACAAAGGCTTAGATTGCTGCTCAAAATTAAGGAATTTTGTAGT 3120  
QY 3135 GGTTCAAAAAATAATTCAACAAGAAACAATACAAAAAAGTGGTAGAATTAACCTATCAC 3194  
| | | | |  
Db 3121 GGTTCAAAAAATAATTCAACAAGAAACAATACAAAAAAGTGGTAGAATTAACCTATCAC 3180  
QY 3195 ATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTAGTATGAGAGTTAGCACTTGAT 3254  
| | | | |  
Db 3181 ATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTAGTATGAGAGTTAGCACTTGAT 3240  
QY 3255 TGAAGATTTCTTTTAAATACTATCAGTTAAACATTTTAATATGATTAATGTATTC 3314  
| | | | |  
Db 3241 TGAAGATTTCTTTTAAATACTATCAGTTAAACATTTTAATATGATTAATGTATTC 3300  
QY 3315 ATTAATGCTACAGAACTGACATTAAGAAATCAATAAATGATTTGTTTACTCT 3364  
| | | | |  
Db 3301 ATTAATGCTACAGAACTGACATTAAGAAATCAATAAATGATTTGTTTACTCT 3350

RESULT 10  
AY017378 3373 bp mRNA linear PRI 28-NOV-2003  
LOCUS Homo sapiens RNA helicase-DEAD box protein RH16 mRNA, complete  
DEFINITION cds.  
ACCESSION AY017378 GI:12621065  
VERSION AY017378.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 3373)  
AUTHORS Cocude,C., Truong,M.-J., Billaut-Mulot,O., Delsart,V.,  
Darcissac,E., Capron,A., Mouton,Y. and Bahr,G.M.  
TITLE A novel cellular RNA helicase, RH16, differentially regulates cell  
growth, programmed cell death and human immunodeficiency virus type  
1 replication  
J. Gen. Virol. 84 (12), 3215-3225 (2003)  
JOURNAL 2 (bases 1 to 3373)  
REFERENCE Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J.,  
AUTHORS Capron,A. and Bahr,G.M.  
TITLE Identification of a new RNA helicase (RH16) regulated by the  
immunomodulator Murabutide  
Unpublished  
3 (bases 1 to 3373)  
JOURNAL Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J.,  
REFERENCE Capron,A. and Bahr,G.M.  
AUTHORS TITLE Direct Submission  
JOURNAL Submitted (09-JAN-2001) Laboratoire d'Immunologie Molculaire de  
1 l'infection et de l'inflammation, Institut Pasteur de Lille, 1 rue  
du Professeur Calmette BP 245, Lille 59019, France  
FEATURES  
source location/Qualifiers  
1..3373  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
155..3232  
/codon\_start=1  
/product="RNA helicase-DEAD box protein RH16"  
/protein\_id="AAG54076.1"  
/db\_xref="GI:12621066"  
/translation="MSGYSTDENFRYLISCFRARKMYIQVEPVLDTLPAPAEVKE  
QIQTIVATSGNQAVELLSTLEKGVHLGWTREVEALRRTGSPLAARYMPELTDL  
PSPFENAHDEYQLNLQPTLVKLVKLDKMEBELLTEDRNRIAAENNGN  
ESGVRELKRIVOKENMFSAFLNVLROGTNNELVQELTGSDCSNAEINLSQVDGP  
QVEBQLSTTVQPNLEKEVGMENSSSESPADSSVSESDTSLAEGSVSLDESIGH  
NSNMGSDSGTMSDSDSEENVAAASPEPEQLRPYOMEVAQPALEGKNIICLPDTS  
KTRVAVYIAKDLDKKKASEPGKIVLVNKKLVLEQLFRKEPQPLKKMYRVIIGLSG  
DTQLKISFEVVKSCDIIISTAQILNSLNLENGEDAGVQLSDSFIIIDECHTNK  
EAVYNNIMRHYLMQKLKNRLKCKENKPIPLPQILGLTASPGVGATKQAKAEHILK  
LCANLDAFTIKYKENLDQLKNQIQEPCKFAIADATREDPFKEKLEIMTRIQTCC  
MSPMSDFGTOPVEQWAIQMEKKAKEGNRESVCAEHLRKYNKALQINDTIRMIDAYT  
HLETFYNEEKDKKRAVIEDSDDEGDEYCDGDEDEDLKKPLKLDETDRFLMTLFE  
NNKMLKRLAENPEYENKLTKLRTIMEQYTRTEESARGIIFTKROSAYALQOWITE  
NEKFAEVGKAAHLLIAGHSSSEFKPMTQNEQKEVISKFRGKINLLIATTVAEEGLDI  
KECNIVIRYGLVTNELAMVQARGARADESTYVLVAHSGSGVIERETVNDPREKMYK  
AIHCVQNMKPEEYAAKILELQMSIMEKMKTKRNIAKHYNKNSLITPLCKNCSVLA  
CSGEDIHVIEKMHVMTPFEKELYIVRENKALQKCADYQINGEIIICKGQAMGTMM  
VHKGLDPLCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCLFSDSD"  
misc\_feature 1148..1201  
/note="Region: DEAD-like helicase domain"  
misc\_feature 1481..1627  
/note="Region: DEAD/DEAH box"  
misc\_feature 2606..2647  
/note="conserved C-terminal helicase region; Region:  
helicase C"  
ORIGIN  
Query Match 99.2%; Score 3338.8; DB 8; Length 3373;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3343; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 15 GAGCCCTGTGACCAACCTCGTCATTTGTCAGGCACAGAGCCGTTAGACCTGCTTCTAAG 74  
Db 1 GGGCCCTGTGACCAACCTCGTCATTTGTCAGGCACAGAGCCGTTAGACCTGCTTCTAAG 60  
QY 75 TGGGCAGCGGACAGCGGCACGCACATTTCACTCTGTCGCCGACAGACAAGACCATCTGCT 134  
Db 61 TGGGCAGCGGACAGCGGCACGCACATTTCACTCTGTCGCCGACAGACAAGACCATCTGCT 120  
QY 135 TGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGTATTCCACAGACGA 194  
Db 121 TGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGTATTCCACAGACGA 180  
QY 195 GAATTTCCGCTATCTCATCTCGTCTTCAGAGGCCAGGGTGAAATGTACATCCAGGTGA 254  
Db 181 GAATTTCCGCTATCTCATCTCGTCTTCAGAGGCCAGGGTGAAATGTACATCCAGGTGA 240  
QY 255 GCCTGTGCTGACCTACCTGACCTTCTGCTGCAGAGGTGAAGAGCAGATTCAAGAGAC 314  
Db 241 GCCTGTGCTGACCTACCTGACCTTCTGCTGCAGAGGTGAAGAGCAGATTCAAGAGAC 300  
QY 315 AGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGTGAGCACCCTTGAGAGAGG 374  
Db 301 AGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGTGAGCACCCTTGAGAGAGG 360  
QY 375 AGTCGGACCTTGTTGACTCGGAAATCTGTGAGAGCCCTCCGGAAACCGGACGCC 434  
Db 361 AGTCGGACCTTGTTGACTCGGAAATCTGTGAGAGCCCTCCGGAAACCGGACGCC 420  
QY 435 TCTGGCCGCCGCTACATGAACCCCTGAGCTCACGGACTTGCCCTTCATCGTTTGAGAA 494  
Db 421 TCTGGCCGCCGCTACATGAACCCCTGAGCTCACGGACTTGCCCTTCATCGTTTGAGAA 480  
QY 495 CGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTACGCCCACTCTGTGTGACAAGCT 554  
Db 481 CGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTACGCCCACTCTGTGTGACAAGCT 540  
QY 555 TCTAGTTAGAGACGCTCTTGATTAAGTGCATGAGAGAGGAACCTGTTGACAAATTGAAGAC 614  
Db 541 TCTAGTTAGAGACGCTCTTGATTAAGTGCATGAGAGAGGAACCTGTTGACAAATTGAAGAC 600  
QY 615 AAACCGGATTGCTGTCGAGAAAACAATGAAATGAATCAGGTGTGAAGAGAGCTACTAA 674  
Db 601 AAACCGGATTGCTGTCGAGAAAACAATGAAATGAATCAGGTGTGAAGAGAGCTACTAA 660  
QY 675 AAGGATTGTGAGAAAGAACTGTTCTGCAATTTCTGAATGTTCTTCTGTCAAACAGG 734  
Db 661 AAGGATTGTGAGAAAGAACTGTTCTGCAATTTCTGAATGTTCTTCTGTCAAACAGG 720  
QY 735 AAACAATGAACCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGAT 794  
Db 721 AAACAATGAACCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGAT 780  
QY 795 TGAGAAATTTATCACAAGTTGATGTCCTCAAGTGAAGAGCAACTCTTTCAACACAGT 854  
Db 781 TGAGAAATTTATCACAAGTTGATGTCCTCAAGTGAAGAGCAACTCTTTCAACACAGT 840  
QY 855 TCAGCCAAATCTGAGAGAGAGGTCTGGGGCATGAGAAATACTCATCAGAAATCATCTTT 914  
Db 841 TCAGCCAAATCTGAGAGAGAGGTCTGGGGCATGAGAAATACTCATCAGAAATCATCTTT 900  
QY 915 TGCAGATTCTTCTGATTTCAGAAATCAGACAAAGTTTGACAGAAAGAGTGCAGCTG 974  
Db 901 TGCAGATTCTTCTGATTTCAGAAATCAGACAAAGTTTGACAGAAAGAGTGCAGCTG 960  
QY 975 CTTAGATGAAGTCTTGACATTAACAGCAATGGGCAGTGAATTCAGGCACCATGGGAAG 1034  
Db 961 CTTAGATGAAGTCTTGACATTAACAGCAATGGGCAGTGAATTCAGGCACCATGGGAAG 1020  
QY 1035 TGATTTCAGATGAAGAGATGTGGCAGCAAGAGCATCCCCGAGGCCAGAACTCCAGCTCAG 1094  
Db 1021 TGATTTCAGATGAAGAGATGTGGCAGCAAGAGCATCCCCGAGGCCAGAACTCCAGCTCAG 1080

QY	1095	GCCTTACC	AAATGGA	AGTTGCCC	CAGCCAGC	CCCTTG	GAAGGA	AGAAATAT	CAATCAT	CTGCCT	1154
Db	1081	GCCTTACC	AAATGGA	AGTTGCC	CAGCCAGC	CCCTTG	GAAGGA	AGAAATAT	CAATCAT	CTGCCT	1140
QY	1155	CCCTAC	GAGGAG	TGGAAAA	CCAGAG	TGGCTG	TTTACA	TTGCCA	AGATCA	CTTAGCAA	1214
Db	1141	CCCTAC	GAGGAG	TGGAAAA	CCAGAG	TGGCTG	TTTACA	TTGCCA	AGATCA	CTTAGCAA	1200
QY	1215	GAAGAAAA	AGCATCT	GAGCCT	GAAAAAG	TTATAG	TTCTTG	CAATAA	GGTACT	GTGCTAGT	1274
Db	1201	GAAGAAAA	AGCATCT	GAGCCT	GAAAAAG	TTATAG	TTCTTG	CAATAA	GGTACT	GTGCTAGT	1260
QY	1275	TGAACAG	CTCTTCC	GCAAGAG	TTCCAAC	CAATTTT	TGAAGAA	TGGTAT	CGTGTTA	TTGG	1334
Db	1261	TGAACAG	CTCTTCC	GCAAGAG	TTCCAAC	CAATTTT	TGAAGAA	TGGTAT	CGTGTTA	TTGG	1320
QY	1335	ATTTAAG	TGATAT	ACC	CAACTG	AAAAATAT	CATTTCC	AGAAGT	TGTCAAG	TCCGTGTAT	1394
Db	1321	ATTTAAG	TGATAT	ACC	CAACTG	AAAAATAT	CATTTCC	AGAAGT	TGTCAAG	TCCGTGTAT	1380
QY	1395	TATTATC	AGTACAG	CTCAAAT	CTTGAAAA	CTCCCTCT	TAAACTT	GAAAAAT	GGAGAGA		1454
Db	1381	TATTATC	AGTACAG	CTCAAAT	CTTGAAAA	CTCCCTCT	TAAACTT	GAAAAAT	GGAGAGA		1440
QY	1455	TGCTG	GTTC	AATGTG	CAGACTT	TCCCTCAT	TATCAT	TGATGA	ATGTCAC	CAACCA	1514
Db	1441	TGCTG	GTTC	AATGTG	CAGACTT	TCCCTCAT	TATCAT	TGATGA	ATGTCAC	CAACCA	1500
QY	1515	CAAGAAG	CAGTGT	ATAATA	CATCAT	GAGCATT	TTTGAT	GCAGA	AGTTGA	AAAACAA	1574
Db	1501	CAAGAAG	CAGTGT	ATAATA	CATCAT	GAGCATT	TTTGAT	GCAGA	AGTTGA	AAAACAA	1560
QY	1575	TAGACTC	AAGAAAA	CAAA	CCAGT	ATTTCC	CTTCAG	ATACTG	GGACTTA	CAGC	1634
Db	1561	TAGACTC	AAGAAAA	CAAA	CCAGT	ATTTCC	CTTCAG	ATACTG	GGACTTA	CAGC	1620
QY	1635	TTCACTG	TGTTG	AGGGGCC	CGAAG	CAAGCC	CAAA	AGCTGA	AGAACAC	ATTTTAA	1694
Db	1621	TTCACTG	TGTTG	AGGGGCC	CGAAG	CAAGCC	CAAA	AGCTGA	AGAACAC	ATTTTAA	1680
QY	1695	ATGTGCC	AATCTT	GATGCA	TTTACT	ATTAAAA	CTGTTAA	AGAAAA	CCTTGAT	CAACTGAA	1754
Db	1681	ATGTGCC	AATCTT	GATGCA	TTTACT	ATTAAAA	CTGTTAA	AGAAAA	CCTTGAT	CAACTGAA	1740
QY	1755	AAACCA	AATCAG	GAGCCAT	GC	CAAGAA	GTTTGG	CCATTG	CGACAT	GCACAGAG	1814
Db	1741	AAACCA	AATCAG	GAGCCAT	GC	CAAGAA	GTTTGG	CCATTG	CGACAT	GCACAGAG	1800
QY	1815	ATTTAA	AGAGAA	CTTCTA	GAAATA	TATGACA	AGGATT	CAAACT	TATGTCA	AATGAGTCC	1874
Db	1801	ATTTAA	AGAGAA	CTTCTA	GAAATA	TATGACA	AGGATT	CAAACT	TATGTCA	AATGAGTCC	1860
QY	1875	AATGTCA	GATTTT	GAACTCA	ACCCTAT	GAAACA	ATGGGCC	ATTCA	AATG	AAAAAAAAGC	1934
Db	1861	AATGTCA	GATTTT	GAACTCA	ACCCTAT	GAAACA	ATGGGCC	ATTCA	AATG	AAAAAAAAGC	1920
QY	1935	TGCAAAA	AAAAAG	GAATCG	CAAGAA	CGTGT	TGTGCA	GAACAT	TTGAG	GAGTACATGA	1994
Db	1921	TGCAAAA	AAAAAG	GAATCG	CAAGAA	CGTGT	TGTGCA	GAACAT	TTGAG	GAGTACATGA	1980
QY	1995	GGCCCT	ACAAATT	AATGACA	CAATTC	GAATG	ATAGATG	CGTAT	CTCAT	CTTGAACTTT	2054
Db	1981	GGCCCT	ACAAATT	AATGACA	CAATTC	GAATG	ATAGATG	CGTAT	CTCAT	CTTGAACTTT	2040
QY	2055	CTATAT	GAGAGAA	AGATTA	GAAAGT	TTGCAG	TCATAGA	AGATG	ATAGT	AGGAGTGG	2114
Db	2041	CTATAT	GAGAGAA	AGATTA	GAAAGT	TTGCAG	TCATAGA	AGATG	ATAGT	AGGAGTGG	2100
QY	2115	TGATGAT	GAGTAT	TGTGAT	GC	TGATGA	AGATGA	AGATG	ATTTAA	AGAAACCTT	2174
Db	2101	TGATGAT	GAGTAT	TGTGAT	GC	TGATGA	AGATGA	AGATG	ATTTAA	AGAAACCTT	2160
QY	2175	GGATG	AAAA	CAGATAG	ATTTCT	CATGACT	TATATTTT	TGAAAA	CAAT	TAATGTTG	2234

Db	2161	GGATGAACAGATAGATTCTCATGACTTTATTTTGGAAAACAATAAAATGTTGAAAG	2220
Qy	2235	GCTGGCTGAAAAACCCAGAAATATGAAAATGAAAAGCTGACCAATTTAAGAAATACCAATAAT	2294
Db	2221	GCTGGCTGAAAAACCCAGAAATATGAAAATGAAAAGCTGACCAATTTAAGAAATACCAATAAT	2280
Qy	2295	GGAGCAATATACTAGAGCTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACA	2354
Db	2281	GGAGCAATATACTAGAGCTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACA	2340
Qy	2355	GAGTGCATATGCGCTTCCAGTGGATTACTGAAAATGAAAAATTGCTGAAGTAGAGT	2414
Db	2341	GAGTGCATATGCGCTTCCAGTGGATTACTGAAAATGAAAAATTGCTGAAGTAGAGT	2400
Qy	2415	CAAAAGCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCATGACACAGAA	2474
Db	2401	CAAAAGCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCATGACACAGAA	2460
Qy	2475	TGAACAAAAAGAAAGTCATTAGTAATTCGACCTGAAAAAATCAATCTGCTTATCGCTAC	2534
Db	2461	TGAACAAAAAGAAAGTCATTAGTAATTCGACCTGAAAAAATCAATCTGCTTATCGCTAC	2520
Qy	2535	CACAGTGCAGAAAGAGTCTGATATTAAAGATGTAAACATTGTTATCCGTTATGCTCT	2594
Db	2521	CACAGTGCAGAAAGAGTCTGATATTAAAGATGTAAACATTGTTATCCGTTATGCTCT	2580
Qy	2595	CGTCACCAATGAATAAGCCATGCTCCAGGCCCGTGTGAGCCAGAGCTGATGACGAC	2654
Db	2581	CGTCACCAATGAATAAGCCATGCTCCAGGCCCGTGTGAGCCAGAGCTGATGACGAC	2640
Qy	2655	CTACGTCCTGGTGTCTCACAGTGGTTCAGAGTTATCGAACATGAGACAGTTAATGATTT	2714
Db	2641	CTACGTCCTGGTGTCTCACAGTGGTTCAGAGTTATCGAACATGAGACAGTTAATGATTT	2700
Qy	2715	CCGAGAGAAGATGATGTATAAAGCTATACATTGTTCAAAATATGAAACAGAGAGTA	2774
Db	2701	CCGAGAGAAGATGATGTATAAAGCTATACATTGTTCAAAATATGAAACAGAGAGTA	2760
Qy	2775	TGCTCATAGATTGGGAATTACAGATGCAAAAGTAAATGAAAAAGAAATGAAACC	2834
Db	2761	TGCTCATAGATTGGGAATTACAGATGCAAAAGTAAATGAAAAAGAAATGAAACC	2820
Qy	2835	GAGAATATATGCAAGCATTTACAGAAATTAACCATCACTAATACTTCTTGCAAAAA	2894
Db	2821	GAGAATATATGCAAGCATTTACAGAAATTAACCATCACTAATACTTCTTGCAAAAA	2880
Qy	2895	CTGCAGTGTCTAGCCTGTTCTGGGGAAGATATCCATGTAAATGAGAAAAATGCATCAGT	2954
Db	2881	CTGCAGTGTCTAGCCTGTTCTGGGGAAGATATCCATGTAAATGAGAAAAATGCATCAGT	2940
Qy	2955	CAATATGACCCCGAATTTCAAGGAATTTACATTGTAAAGAAAAAACAAGCACTGCAAAA	3014
Db	2941	CAATATGACCCCGAATTTCAAGGAATTTACATTGTAAAGAAAAAACAAGCACTGCAAAA	3000
Qy	3015	GAAGTGTGCCGACTATCAAAATAAATGTGAAATCATCTGCAAAATGTGGCCAGGCTGGGG	3074
Db	3001	GAAGTGTGCCGACTATCAAAATAAATGTGAAATCATCTGCAAAATGTGGCCAGGCTGGGG	3060
Qy	3075	AACAATGATGGTGACAAGAGGCTTAGATTTGCTGTCTCAAAATAAGGAATTTGTAGT	3134
Db	3061	AACAATGATGGTGACAAGAGGCTTAGATTTGCTGTCTCAAAATAAGGAATTTGTAGT	3120
Qy	3135	GGTTTCAAAAAATTAATTCACAAAGAAACAATACAAAAAAGTGGGTAGATTACCTATCAC	3194
Db	3121	GGTTTCAAAAAATTAATTCACAAAGAAACAATACAAAAAAGTGGGTAGATTACCTATCAC	3180
Qy	3195	ATTTCCCAATCTGACTAATTCAGAAATGCTGTTAATTTAGTGATGAGGATTAAGCACTGAT	3254
Db	3181	ATTTCCCAATCTGACTAATTCAGAAATGCTGTTAATTTAGTGATGAGGATTAAGCACTGAT	3240
Qy	3255	TGAAGATTCTTTAAATACTATCAGTTAAACATTTAATATGATTATGATTATGTATTC	3314

Db 3241 TGAAGATTCTTTAAATACTATCAGTTAAACATTTAATATGATTAATGATTAATGATATTC 3300

Qy 3315 ATTATGCTACAGACTGCATAGAATCAATAAATGATTTGTTACTCT 3364

Db 3301 ATTATGCTACAGACTGCATAGAATCAATAAATGATTTGTTACTCT 3350

RESULT 11

BC078180 2540 bp mRNA linear PRI 27-JUL-2004

LOCUS BC078180

DEFINITION Homo sapiens interferon induced with helicase C domain 1, mRNA (cDNA clone IMAGE:6502757), partial cds.

ACCESSION BC078180

VERSION BC078180.1 GI:50415794

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2540)

REFERENCE AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Abramson,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mulahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Sanchez,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 2540)

REFERENCE AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (19-JUL-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

REMARK COMMENT Contact: MGC help desk

REMARK COMMENT Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)

REMARK COMMENT Tissue Procurement: ATCC

REMARK COMMENT cDNA Library Preparation: Life Technologies, Inc.

REMARK COMMENT cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

REMARK COMMENT DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),

REMARK COMMENT Gaithersburg, Maryland;

REMARK COMMENT Web site: <http://www.nisc.nih.gov/>

REMARK COMMENT Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)

REMARK COMMENT Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Grante,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsugeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Series: IRAK Plate: 174 Row: p Column: 16

FEATURES

source

1..2540

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6502757"

/tissue\_type="uterus, leiomyosarcoma"

/clone\_1b="NIH MGC\_71"

/lab\_host="DH10B"

/note="vector: pCMV-SPORT6"

1..>2540

/gene="IFIH1"

/note="synonyms: MDA5, MDA-5, H1cd"

/db\_xref="GeneID:64135"

/db\_xref="MIM:606951"

224..>2540

/gene="IFIH1"

/codon\_start=1

/product="IFIH1 protein"

/protein\_id="AAH78180.1"

/db\_xref="GI:50415795"

/db\_xref="GeneID:64135"

/db\_xref="MIM:606951"

/translation="MSGYSTDENFRYLISCFRARKYIQLPEVLDTPLPAEVE QIORIVATSGMNAVELLSLEKGVHMGWTRFVALRRTGSPLAARYMNEELDL PPSPFENAHDEYLQNLNLOPLVDKLVRDVLDKMEELLTIEDNRNIAAENGN ESGVRELKRIVQKENWFSAPLVLRQTGNNEVLQELTSGDCSESNAEIENSQYDGP OVEBOLSTVQPNLEKEVWGNENSSSESFADSSVSESDTSLAEGSVSCDESLGH NSNMGSDSGTMSDSEBNVAPASPEPELQRPYQMEVAQPALEKNITICLPTSGG KTRVAVYIAKDHLDKKKKASEGKVIIVNVKVLVEQLPRKEQPLKKWRYIGLSG DTQKISPEPEVVKSCDIISTAQILENSLNLNLENGEDAGVQSDPSLIIDECHTNK EAVYNNIMRHYLMQKLNKNRKKENKPVIPPLPOLGLTASPGVGATKQAKAEHILK LCANLDAFTIKTKENLDQKNQIQEPCKKFAIADATREDPFKEKLEIMTRIQTYCQ MSPMSDGTQPYEQWAIQMEKKAKEGNRKERVCAEHARKNEALQINDTIMDAYT HLETFYNEBKDKKFAVIEDSDGDDDEYCDGEDBDLKKPLKLDETDRFLMTLFEENKMLKRLAENPEYENKLTALNTIMEQYTRTEESARGIIFTKTRQSAVALSQWITE NEKFAEVGVKAHHLIGAGHSSEFKPMTOKKKK"

ORIGIN

Query Match 73.6%; Score 2476; DB 8; Length 2540;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2479; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCGCGCCGCGCTGAGACCCCTGTGACAACTCGTCATTGTACAGCACAGAGCGGTAGAC 60

Db 56 GCGCGCGCGCTGAGACCCCTGTGACAACTCGTCATTGTACAGCACAGAGCGGTAGAC 115

Qy 61 CCGCTTCTTAAGTGGGCGAGCGGACGACGACATTTCACTGTCGCCGACAGCAA 120

Db 116 CCGCTTCTTAAGTGGGCGAGCGGACGACGACATTTCACTGTCGCCGACAGCAA 175

Qy 121 CAGCACCATCTGCTGGGAGAACCCCTCCCTTCTTGAGAAAGAAAGATGCGAATGG 180

Db 176 CAGCACCATCTGCTGGGAGAACCCCTCCCTTCTTGAGAAAGAAAGATGCGAATGG 235

Qy 181 TATTCACAGACGAGAAATTCGGCTATCTCATCTCGTGCTTCAGGGCCAGGGTGAATAATG 240

Db 236 TATTCACAGACGAGAAATTCGGCTATCTCATCTCGTGCTTCAGGGCCAGGGTGAATAATG 295

Qy 241 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCCTGCTGACAGAGGTGAAGAG 300

Db 296 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCCTGCTGACAGAGGTGAAGAG 355

Qy 301 CAGATTGAGAGGACAGTCCGACCTCCGGGAACATGACAGGACATTGAACCTGCTGAGC 360

Db 356 CAGATTGAGAGGACAGTCCGACCTCCGGGAACATGACAGGACATTGAACCTGCTGAGC 415

Qy 361 ACCTTGAGAGGAGTCTGGCACCTTGTTGACTCGGGAATTCGTGAGGCGCTCCGG 420

Db 416 ACCTTGAGAGGAGTCTGGCACCTTGTTGACTCGGGAATTCGTGAGGCGCTCCGG 475

QY 421 AGAACCGGACGCCCTCTGGCCCCCGCTACATGAACCCGTAGCTCACGCACTTGCCCTCT 480  
|||||  
Db 476 AGAACCGGACGCCCTCTGGCCCCCGCTACATGAACCCGTAGCTCACGCACTTGCCCTCT 535  
QY 481 CCATCGTTTGAGAACGCTCATGATGATATCTCAACTGCTGAACCTCTTCAGCCCACT 540  
|||||  
Db 536 CCATCGTTTGAGAACGCTCATGATGATATCTCAACTGCTGAACCTCTTCAGCCCACT 595  
QY 541 CTGGTGACAAAGCTTCTAGTTAGAGACGCTTGATTAAGTGCATGAGAGAACTGTTG 600  
|||||  
Db 596 CTGGTGACAAAGCTTCTAGTTAGAGACGCTTGATTAAGTGCATGAGAGAACTGTTG 655  
QY 601 ACAATTGAGACAGAAACCGGATTGCTGCTGCAGAAACAATGGAATGAATCAGGTGA 660  
|||||  
Db 656 ACAATTGAGACAGAAACCGGATTGCTGCTGCAGAAACAATGGAATGAATCAGGTGA 715  
QY 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAATGCTGCTCTGCAATTCTGAATGT 720  
|||||  
Db 716 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAATGCTGCTCTGCAATTCTGAATGT 775  
QY 721 CTTGCTCAACAGGAAACAATGAATCTTGTCCAGAGTTAACAGGCTCTGATTGCTCAGAA 780  
|||||  
Db 776 CTTGCTCAACAGGAAACAATGAATCTTGTCCAGAGTTAACAGGCTCTGATTGCTCAGAA 835  
QY 781 AGCAATGCAGAGATTGGAATTTATCAACAAGTTGATGCTCCTCAAGTGAAGACAACTT 840  
|||||  
Db 836 AGCAATGCAGAGATTGGAATTTATCAACAAGTTGATGCTCCTCAAGTGAAGACAACTT 895  
QY 841 CTTTCAACCAACAGTTCCAGCCAATCTGAGAGAGAGTCTGGGCGATGAGATAACTCA 900  
|||||  
Db 896 CTTTCAACCAACAGTTCCAGCCAATCTGAGAGAGAGTCTGGGCGATGAGATAACTCA 955  
QY 901 TCAGAAATCATCTTTTGGCAGATTCTTCTGTAGTTTCAGAATCAGACACAAGTTTGGCAGAA 960  
|||||  
Db 956 TCAGAAATCATCTTTTGGCAGATTCTTCTGTAGTTTCAGAATCAGACACAAGTTTGGCAGAA 1015  
QY 961 GGAAGTGTCAAGCTGTTAGATGAAGTCTTGACATTAACAGCAACATGGGCACTGATTCA 1020  
|||||  
Db 1016 GGAAGTGTCAAGCTGTTAGATGAAGTCTTGACATTAACAGCAACATGGGCACTGATTCA 1075  
QY 1021 GGCAACCATGGGAAGTATCAGATGAAGAAATGTGGCAGCAAGACATCCCCGAGCCA 1080  
|||||  
Db 1076 GGCAACCATGGGAAGTATCAGATGAAGAAATGTGGCAGCAAGACATCCCCGAGCCA 1135  
QY 1081 GAACTCCAGCTCAGGCTTAACCAATGGAAGTTGCCAGCCAGCTTGAAGGGAAGAT 1140  
|||||  
Db 1136 GAACTCCAGCTCAGGCTTAACCAATGGAAGTTGCCAGCCAGCTTGAAGGGAAGAT 1195  
QY 1141 ATCATCATCTGCTCCCTACAGGAGTGAAGAAACCAGAGTGGCTGTTTACATTGCCAAG 1200  
|||||  
Db 1196 ATCATCATCTGCTCCCTACAGGAGTGAAGAAACCAGAGTGGCTGTTTACATTGCCAAG 1255  
QY 1201 GATCACTTAGACAAGAAAAAAGCATCTGAGCTGAAAAAGTTATAGTTCTTGTCAAT 1260  
|||||  
Db 1256 GATCACTTAGACAAGAAAAAAGCATCTGAGCTGAAAAAGTTATAGTTCTTGTCAAT 1315  
QY 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAAACCATTTTGAAGAAATGG 1320  
|||||  
Db 1316 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAAACCATTTTGAAGAAATGG 1375  
QY 1321 TATCGTGTATTGATTAAGTGTGATACCAACTGAAATATATCATTTCCAGAAGTTGTC 1380  
|||||  
Db 1376 TATCGTGTATTGATTAAGTGTGATACCAACTGAAATATATCATTTCCAGAAGTTGTC 1435  
QY 1381 AAGTCTGTGATATTATATCATGACAGCTCAAAATCTTGAAAACTCCCTTTAAACTTG 1440  
|||||  
Db 1436 AAGTCTGTGATATTATATCATGACAGCTCAAAATCTTGAAAACTCCCTTTAAACTTG 1495  
QY 1441 GAAAAATGAGAGAGATGCTGCTGTTCAATTGTTCAGACTTTTCCCTCATTAATCATGTGAA 1500  
|||||  
Db 1496 GAAAAATGAGAGAGATGCTGCTGTTCAATTGTTCAGACTTTTCCCTCATTAATCATGTGAA 1555  
QY 1501 TGTTCATCACACCAACAAAGAGCAGTGTATATAATCATCATGAGCGCATTAATTGATGCAG 1560  
|||||

Db 1556 TGTTCATCACACCAACAAAGAGCAGTGTATTAATTAACATCATGAGCGATTATTGATGCAG 1615  
QY 1561 AAGTGAACCAATAGACTCAAGAAAGAAACAAACAGATGATTCCTCTCAGATA 1620  
|||||  
Db 1616 AAGTGAACCAATAGACTCAAGAAAGAAACAAACAGATGATTCCTCTCAGATA 1675  
QY 1621 CTGGGACTAACAGCTTCACTGCTGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
|||||  
Db 1676 CTGGGACTAACAGCTTCACTGCTGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1735  
QY 1681 CACATTTTAAACTATGTGCCAATCTTGATGCAATTTACTATTAAACTGTAAAGAAAC 1740  
|||||  
Db 1736 CACATTTTAAACTATGTGCCAATCTTGATGCAATTTACTATTAAACTGTAAAGAAAC 1795  
QY 1741 CTTGATCAACTGAAAAACCAATACAGAGGCCATGCAAGAGTTTGGCAATTGCAAGATGCA 1800  
|||||  
Db 1796 CTTGATCAACTGAAAAACCAATACAGAGGCCATGCAAGAGTTTGGCAATTGCAAGATGCA 1855  
QY 1801 ACCAGAGAGATCCATTTAAAGAGAACTTCTAGAATTAATGACAAGATTCAAACTTAT 1860  
|||||  
Db 1856 ACCAGAGAGATCCATTTAAAGAGAACTTCTAGAATTAATGACAAGATTCAAACTTAT 1915  
QY 1861 TGTCAATGAGTCCATGTCAAGATTTTGAACCTCAACCTATGAACAATGGCCATTCAA 1920  
|||||  
Db 1916 TGTCAATGAGTCCATGTCAAGATTTTGAACCTCAACCTATGAACAATGGCCATTCAA 1975  
QY 1921 ATGGAACCAAAAGCTGCAACCAAAAGAAATCCCAAAAGACGTGTTGTGCAAGACATTTG 1980  
|||||  
Db 1976 ATGGAACCAAAAGCTGCAACCAAAAGAAATCCCAAAAGACGTGTTGTGCAAGACATTTG 2035  
QY 1981 AGGAATGACAAATGAGGCCCTCAAAATTAATGACACAATTCGAATGATAGATGCGTATACT 2040  
|||||  
Db 2036 AGGAATGACAAATGAGGCCCTCAAAATTAATGACACAATTCGAATGATAGATGCGTATACT 2095  
QY 2041 CATCTTGAACCTTCTATAATGAAGAGAAAGATAAGAGTTTGCAGTATAGAAGATGAT 2100  
|||||  
Db 2096 CATCTTGAACCTTCTATAATGAAGAGAAAGATAAGAGTTTGCAGTATAGAAGATGAT 2155  
QY 2101 AGTGATGAGGCTGTGATGATGATGATTTGTGATGTGATGAAGATGATGATTTAAAG 2160  
|||||  
Db 2156 AGTGATGAGGCTGTGATGATGATGATTTGTGATGTGATGAAGATGATGATTTAAAG 2215  
QY 2161 AAACCTTTGAAACTGTGATGAACAGATAGATTCTCATGACTTTATTTTGAACCAAT 2220  
|||||  
Db 2216 AAACCTTTGAAACTGTGATGAACAGATAGATTCTCATGACTTTATTTTGAACCAAT 2275  
QY 2221 AAAATGTTGAAAAAGCTGCTGAAAAACCAAGATATGAAAAATGAAAAGCTGACCAATTA 2280  
|||||  
Db 2276 AAAATGTTGAAAAAGCTGCTGAAAAACCAAGATATGAAAAATGAAAAGCTGACCAATTA 2335  
QY 2281 AGAAATACCATTAATGAGCAATATATCTAGACTGAGGAATCAGCAGAGAAATATCTTT 2340  
|||||  
Db 2336 AGAAATACCATTAATGAGCAATATATCTAGACTGAGGAATCAGCAGAGAAATATCTTT 2395  
QY 2341 ACAAAAAACAGCAGACAGTGCATATATGCGCTTCCAGTGAGTTACTGAAAAATGAAAAATTT 2400  
|||||  
Db 2396 ACAAAAAACAGCAGACAGTGCATATATGCGCTTCCAGTGAGTTACTGAAAAATGAAAAATTT 2455  
QY 2401 GCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGAGAGCTGGAACAGCAGAGTGAAT 2460  
|||||  
Db 2456 GCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGAGAGCTGGAACAGCAGAGTGAAT 2515  
QY 2461 CCCATGACACAGAAATGAACAAAAA 2484  
|||||  
Db 2516 CCCATGACACAGAAATGAACAAAAA 2539  
|||||

RESULT 12  
CO777379 3771 bp DNA linear PAT 11-MAR-2004  
LOCUS CO777379  
DEFINITION Sequence 1065 from Patent EP1394274.  
ACCESSION CO777379

VERSION	CQ777379.1	GI:45380363
KEYWORDS	Mus musculus (house mouse)	
SOURCE	Mus musculus	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.	
REFERENCE	1 Ohtani, N., Sugita, Y., Yamaya, M., Kubo, H., Nagai, H. and Izuhara, K. Methods of testing for bronchial asthma or chronic obstructive pulmonary disease Patent: EP 1394274-A 1065 03-MAR-2004;	
JOURNAL	Genex Research, Inc. (JP)	
FEATURES	Location/Qualifiers	
source	1..3771	
	/organism="Mus musculus"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:10090"	
ORIGIN		
Query Match	65.5%;	Score 2202.4; DB 6; Length 3771;
Best Local Similarity	80.7%;	Pred. No. 0;
Matches 2609;	Conservative 0;	Mismatches 616; Indels 7; Gaps 3
QY	121 CAGCACCATCTGCTTGGAGAACCTCTCCC-TTCTTGAGAAAGAAATGTCGAATGG	179
DB	186 CCGCCCCCGCCCGGAGACTCTCTCCATTCTCTGAGACCTCAGATGTGATTGT	245
QY	180 GTATTCCACAGACGAGAATTTCGGTATCTCATCTCGTTCAGGCCAGGGTGAAT	239
DB	246 CTGTCTGACAGAGACAGCTTACGAATCTCATCTTATTCTTACGCCAGCTGAAT	305
QY	240 GTACATCCAGGTGAGCGCTGTGCTGAACCTGACCTTCTGCTGACAGGTGAAGA	299
DB	306 GTACATTCAGGTGAGCGCAGTGTGACCACTCATCTTCTGTGACAAACCAAGA	365
QY	300 GCAGATTACAGAGACAGTCCCACTTCGGGAAACATGACGACGTTGAATCTGCTGAG	359
DB	366 GCAGATTCTTAAAGATCAACACCTGTGTAAACACGCGCGGACAGACTGCTGAG	425
QY	360 CACCTTGAGAAAGGAGTCTGCACTTGTGGACTCGGAATTCGTGAGGCCCTCCG	419
DB	426 CACCTTGAGAGCAGGACATGGCTCTGGAGTGAACGAGATGTTCTGAGAGCCCTAGA	485
QY	420 GAGAACCGGCAGCCCTCTGGCCCGCCGCTACATGAACCTGAGCTCAGGACTGCCCC	479
DB	486 GCACAGTGGCAATCCCTAGCCCGCGCTATGTCAAAACCACTCAGTATCTGCCCTC	545
QY	480 TCATCTGTTGAGAACGCTCATGTATATCTCCAACCTGCTGAACCTCTTCAGCCCCAC	539
DB	546 TCCCTCTCTGAGACTGCCCATGACGAGTGTCTCACTTGTGACCTCTCCAGCCAC	605
QY	540 TCTGTGACAAAGCTTCTAGTTAGAGCGTCTGATAGTGCATGAGAGGAACTGTT	599
DB	606 TTTGTGAGACAACTTCTGATTAAGATGCTTGGACACTTGTGAGAGGAACTATT	665
QY	600 GACAAATTGAAGACAGAAACCGGATTGCTGTCAGAAACCAATGAAATGATCAGGTGT	659
DB	666 GACAGTCGAGACAGAAATCGGATTTCTGTGTCAGGAAACAGCGGAAATGATCAGGTGT	725
QY	660 AAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAAGCTGTTCTCTGCAATTCTGAATGT	719
DB	726 AAGAGAGCTGTGAGAGGATTGTGAGAGGAAAGAACTGTTTCTTCACTTCTGATGT	785
QY	720 TCTTCGTCAAAACAGGAAACCAATGAACCTGTTCAGAGATTAAACAGGCTGTGCTCAGA	779
DB	786 TCTGCGCCAACTGGAATGATGACATTCCTCAAGAACTAACAGGTGAGGCTGCCAGA	845
QY	780 AAGCAATGACAGAGATTGAGATTATACAACTTGATGCTCTCAAGTGAAGACAAT	839
DB	846 AGACAACACAGACTTGCTTAATCTGCTACAGAGATGGGCTGACGCTTAATGAGTGTCT	905
QY	840 TCTTTCAACCAAGTTCAGCCCAATCTGAGAGGAGGCTCTGGGGCATGAGAAATAATCTC	899

Db	906	TCTGCCCTGCTGTCGATGAGTCAAGTCTGGAGACAGAGGCTGGAACTGACGACATATT	965
Qy	900	ATCAGAATCATCTTTGGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAACTTTGGCAGA	959
Db	966	ACCAAGAGGCTTCTGTACAGATTGCTGTGACCAAGAATCAGACACAACTTTGGCAGA	1025
Qy	960	AGGAAGTGTACGTGCTTAGATGAAAGTCTTGGACATTAACAGCAATGGGCACTGATT	1019
Db	1026	AGGAAGTGTACGTGCTTAGATGAAAGTCTTGGACATTAACAGCAATGGGCACTGATT	1085
Qy	1020	AGGCACCATGGGAAGTGAATTCAGATGAA--GAGAAATGTGGCAGCAAGAGCATCCCGGA	1076
Db	1086	AGGCACCATGGGAAGTGAATTCAGATGAAAGTGTCAACAGACAAAAGAGTATCCCGGA	1145
Qy	1077	GCCAGAACTCCAGCTCAGGCTTACCAAAATGGAAGTGGCCAGCCAGCCTTGGAAAGGAA	1136
Db	1146	GCCAGAACTCCAGCTCAGGCTTACCAAAATGGAAGTGGCCCAACAGCTCTAGATGGAA	1205
Qy	1137	GAATATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTGC	1196
Db	1206	GAATATTATTATCTGCTCCCAACGGGAGTGGAAAAACAGAGTGGCTGTTTACATTAC	1265
Qy	1197	CAAGGATCATTAGACAGAAGAAAAAGCATCTGAGCCTGGAAAAATTATAGTCTTGT	1256
Db	1266	CAAGATCATTAGACAGAAGAAAGCAGGATCTGAATCCGGGAAGTTATCTTCTGT	1325
Qy	1257	CAATAAGTACTGTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACCATTTTGAAGAA	1316
Db	1326	CAATAAGTATGTGTAGCAAGAACACTTTCCGAAAAAGATTCAACCCATATTGAAGAA	1385
Qy	1317	ATGCTATCGTGTATTGGAATTAAGTGTGATACCCTGAATAATATCATTTCCAGAAGT	1376
Db	1386	ATGCTATCGAATTATTGATTAAGTGGCGATACCCTGAATAATATCATTTCCAGAAGT	1445
Qy	1377	TGTCAGTCTCTGTGATTTATTATCAGTACAGCTCAAACTCTGAAAACTCCCTTTAA	1436
Db	1446	TGTCAAATCTTACGATGTATTATCAGACACTGCTCAAACTCTGAAAACTCCCTTTAA	1505
Qy	1437	CTTGAAAAATGAGAGATGCTGTGTTCAATTGTCAGACTTTTCCCTCATATTATTTGA	1496
Db	1506	TCTGAGAGTGGAGCGATGACGGTGTGACGCTGTCAAGCTTCTCTCATTTATCATTTGA	1565
Qy	1497	TGAATGTCATCACACCAACAAAGAGCAGTATTAATAATCATCATGAGGCATTTTGAT	1556
Db	1566	TGAGTCCATCACACCAACAAAGAGGCAGTCTATTAACAATCATGAGCGATTTTGAA	1625
Qy	1557	GCAGAAAGTGAACCAATAGACTCAAGAAAGAAAAACAACAGTATCCCTTCTCTCA	1616
Db	1626	GCAGAAAGTGAACCAATAGACTCAAGAAAGAAAAACAACAGTATCCCTTCTCTCA	1685
Qy	1617	GATACTGGACTAACAGCTTCACTGTGTGTTGAGGGGCCACGAAGCAAGCCAAAGCTGA	1676
Db	1686	GATACTAGGACTGACAGCTTCACTGTGTGTTGAGAGCCAAAAAGCAGTCTGAGGCTGA	1745
Qy	1677	AGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACTGTTAAAGA	1736
Db	1746	AAAAACATATTTAAATATATGTGCCAATCTTGATGCTTTTACCATTTAAAAACAGTGAAGA	1805
Qy	1737	AAACCTTGATCAACTGAAAAACCAAAATACAGAGGCATGCAAGAAAGTTGCCATTGAGA	1796
Db	1806	GAATCTTGTCAACTCAAAACCAAAATTAAGAACATGCAAGAAATTTGTGATTGCTGA	1865
Qy	1797	TGCAACCAAGAGAAGATCCATTTAAAGAGAACTTCTAGAAATATGACAAAGATTCAAA	1856
Db	1866	TGACACCAAGAGAAATCCATTTAAAGAGAACTTCTAGAAATATGACAAAGATTCAAG	1925
Qy	1857	TTATTTGCAAAATGAGTCCAATGTCAAGTTTGGAACTCAACCCATATGAACAATGGCCAT	1916
Db	1926	TTACTGCCAAAAAGTCCAATGTCAAGTTTGGAACTCAACCCATATATGAGCAGTGGCCAT	1985
Qy	1917	TCAAAATGAAAAAAAGCTGCAAAAAAGGAATTCGCAAGAAAGCTGTTGTGCAAGACA	1976

Db	1986	TCAAATGGAGAAAAAAGCTGCTTAAAGACGGAAATCGCAAGAATCGCGCTGTGTGCAGAAC	2043
Qy	1977	TTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACAAATTCGAATGATAGATCGTA	2036
Db	2046	TTTGAGGAAGTACAACGAAGCCCTCAAAATCAACGACACGATCCGAATGATTGATGATA	2105
Qy	2037	TACTCATCTGAAACTTCTATAATGAGAAGAAAGATAAGAGTTTGCACTCATAGAA	2096
Db	2106	TAGCCACCTGAGGCACTTCTACACTGATGAGAAAGAAAGAGTTCCGAGTCCT--CA	2162
Qy	2097	TGATAGTATGAGGGGTGCTGATGATGAGTATTGTGATGGTGAATGAAGATGAGATGATT	2156
Db	2163	TGACAGCGACAAGAGTATGACGAGGCCAGCAGTTGCAATGACCAACTTAAGGCGATGT	2222
Qy	2157	AAAGAAACCTTTGAAACTGATGAAACAGATAGATTCTCATGACTTTATTTTGGAAA	2216
Db	2223	AAAGAAATCTTTGAAACTGACGAAACGATGAATTTCTCATGAATTTGTTCTTGTATA	2282
Qy	2217	CAATAAATGTTGAAAAAGCTGCTGAAAAACCAAGATATGAAAAATGAAAAGCTGACCA	2276
Db	2283	CAAGAAATGTTGAAAAAACTAGCTGAAAAACCAAAATACGAGAATGAAAACTCATTA	2342
Qy	2277	ATTAAGAAATACCATTAATGAGCAATATACTAGACTGAGAAATCAGCAGAGGAATAT	2336
Db	2343	ATTAAGAAACACGATACTGGAACAAATTCACAAAGGTCTGAGSAGTCTCTCCGAGGAATTAT	2402
Qy	2337	CTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAA	2396
Db	2403	TTTTACAAAAACACGACAGACACTACGCACTTCCAGTGGATCATGAAAAATGCAA	2462
Qy	2397	ATTTGCTGAAGTAGAGTCAAGCCCAACCATCTGATTGGAGCTGACACAGCAGTAGTT	2456
Db	2463	GTTTGGGAAGTTGGAGTCAAGCGCATCACCTGATTGGCGCGGGGCACAGCAGTAGT	2522
Qy	2457	CAAAACCATGACACAGATGAACAAAAAGAGTCATTAGTAATTTCCCACTGAAAAAT	2516
Db	2523	CAAGCCCATGACTCAGACTGAACAAAAAGAGTCATTAGTAATTTCCCACTGGCGAAAT	2582
Qy	2517	CAATCTGCTTATCGCTACACAGTGGCAGAGAAGGCTTGATATTAAGAATGTAACT	2576
Db	2583	AAATCTGCTTATCGCTACACGCTGGCAGAGAAGGCTTGATATCAAAAGAGTCAATAT	2642
Qy	2577	TGTTATCCGTTATGCTCTCGTCAACCAATGAATAGCCATGCTCCAGGCCGCTGTCAGC	2636
Db	2643	TGTTATTCGTTATGCGCTTGTCACGAACGAGATAGCCATGCTCCAGGCCGCGGCTGAGC	2702
Qy	2637	CAGAGCTGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGACA	2696
Db	2703	CAGAGCTGATGAAGACGTAATGTCCTGTCACCAAGCAGTGGCTCAGAGTTACCGAACG	2762
Qy	2697	TGAGACAGTTAATGATTTCCGAGAGAAGATGATGATTAAGCTATACATTTGTTCAAA	2756
Db	2763	GGAGATTGTTAATGATTTCCGAGAGAAGATGATGATTAAGCTATTAACCGTTCAAA	2822
Qy	2757	TATGAAACCAAGAGATATGCTCATTAAGATTTTGGAAATTCAGATGCCAAAGTATAATGA	2816
Db	2823	CATGAAACCAAGAGATATGCACTAAGATTTTGGAAATTCAGAGTGCAAAAGTATCTTGA	2882
Qy	2817	AAAGAAATGAAAAACCAAGAAATATTGCCAAGCATTACAGAATTAACCCATCACTAAT	2876
Db	2883	AAAGAAATGAAAGTCAAAAGAACATTTGCCAAGCAATACAGCAATCCATCGTTAAT	2942
Qy	2877	AACCTTCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGCGGAAAGATATCCATGTAAT	2936
Db	2943	AACACTTCTTGCAAAAAATGTAGCATGTGCTGCTGCGGAGAAAAACATCCATGTCAAT	3002
Qy	2937	TGAGAAATGCATCAGCTCATATATGACCCCAAGATTCAGGAACCTTACATTGTAAGAGA	2996
Db	3003	TGAGAAAGATGCATCATGTCAATATGACACCAAGATTTCAAGGAGCTCTACATTGTAAGAGA	3062
Qy	2997	AAACAAAGCACTGCAAAAAGAGTGTCCGACTATCAAAATAAAATGCTGAATCATCTGCAA	3056
Db	3063	AAACAAAGCACTGCAAAAAGAAATTTGCTGATTATACAGCAATGAGAGATTATCTGCAA	3122

QY	3057	ATGTGGCCAGGCTTGGGGAACAATGATGGTGACAAAGGCTTAGATTTGCCCTGTCTCAA	3116
Db	3123	GTGTGGCCAGGCTTGGGGAACAATGATGGTGACAAAGGCTTAGATTTGCCCTGTCTTAA	3182
QY	3117	AATAAGGAATTTTGTAGTGGTTTTCAAAAATAATTCACAAAGAAACAATACAAAAAGTG	3176
Db	3183	AATAAGGAATTTTGTAGTCAATTTTCAAAAAATACTCACCGAAGAAACAGTACAGAAGTG	3242
QY	3177	GGTAGAATTACCTATCACATTTTCCCAATCTTGACTATTCAGAATGCTGTTTATTAGTGA	3236
Db	3243	GGTGAATTGCCCTATCAGATTTCTGTACTCTTGACTACTCAGAATACTGCTGTATAGTGA	3302
QY	3237	TGAGGATTAGCACTGATTGAAGATCTTTTAAATACTATCAGTTAAACATTTAATATG	3296
Db	3303	TGAAGATTAGCACTGATTGATGATTATTAAATACTGCCAACTCAACATTTGAATATG	3362
QY	3297	ATTATGATTAAATGATTCATTATGCTACAGACTGACATGAATCAATATAA	3348
Db	3363	ATTTAATTTGTTTTCCTACACTGAGCTAAGCTTGAATATAGTAGA	3414
RESULT 13			
AF374384		3771 bp	linear
LOCUS			ROD 02-MAY-2002
DEFINITION	AF374384	Mus musculus strain C57BL/6J melanoma differentiation associated	
ACCESSION	AF374384	gene 5-like protein mRNA, complete cds.	
VERSION	AF374384.1	GI:20385627	
KEYWORDS			
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;	
		Sciurognathi; Murioidea; Muridae; Murinae; Mus.	
REFERENCE		1 (bases 1 to 3771)	
AUTHORS		Kang, D.-C. and Fisher, P.B.	
TITLE		Murine homolog of melanoma differentiation associated gene-5 - a	
JOURNAL		DEXH group RNA dependent ATPase	
REFERENCE		2 (bases 1 to 3771)	
AUTHORS		Kang, D.-C. and Fisher, P.B.	
TITLE		Direct Submission	
JOURNAL		Submitted (26-APR-2001) Urology, Columbia University, P & S, 630	
		West 168th, New York, NY 10032, USA	
FEATURES		Location/Qualifiers	
source		1..3771	
		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/strain="C57BL/6J"	
		/db_xref="taxon:10090"	
		235..3312	
		/note="RNA-dependent ATPase; predicted molecular weight:	
		115954.50 Da"	
		/codon_start=1	
		/product="melanoma differentiation associated gene 5-like	
		protein"	
		/protein_id="AAM21359.1"	
		/db_xref="GI:20385628"	
		/translation="MSIVCSAEDSFRNLILFRPRLKMYIQVEPVLHILFLSAETKE	
		QILKINTCGNTSAEELLSTLEOGWPLWTOMFEAL EHSNGPLAARYVKPTLTDL	
		PSPSSETAHDECLHLTLLOPTLVKLLINDVLTDFEKGGLTVEDNRISAAGNSGN	
		ESGVRELRLRIVOKENMFSTFLDVARQTNDALFOELTGGCCPEDNTDLANSSHRDGP	
		AANECLLPAVDESSLETEAMVNDILPEASCTDSSVTTESDTSIAEGSVSCFDESLGH	
		NSNMGSDSGTMSDSDESIVITQTRVSEPELQLRPYOMEAQPALDGNIIICLPTGS	
		GKTRVAVYITKDHLDKKQASGKIVLVNKMVLAEOFLRKEFNRYLKKMYRIIGLS	
		GDTOLEKISFEPEVVKSYDVIISTAQILNLSLNLBESGDDGVQLSDPSLIIDECHTNT	
		KEAVYNNIMRRYLKQKLRNNDLKQNKPAIPLPQILGLTASPGVGAARKOSEAEKHIL	
		NICANDAFITIKTVKENLGOQLKHQIKERCKFVIADPTRENPFKEKLEIMASIQTVC	
		OKSPMSDFGTQHYEQWAIOMEKKAODGNRKDRVCAEHLRKVNEALQINDTIRMIDAY	
		SHLEAFYDKEKKFAVLNDSKSDDEASSCNDQLKGDVKSLKLDETDFLMLNLFED	
		NKKMLKLAENPKYENKILKRLNTLLEQFTRSESSRGIIFTKTROSTYALQWIME	
		NAKFAEVGVKAHHLIGAGHSSEVKEMTQTEQKEVISKFRTEGINLLIATTVAEGLDI	



QY	2097	TGATAGTATGAGGGTGGTGA	TGATGAGTATTTGATGGTGA	TGAAGATGAGATGATTT	2156
Db	2163	TGACAGCGACAAGAGTGA	CGAGGCCAGCAGTTGCAATGACCAACTTAAGGCGATGT		2222
QY	2157	AAAGAAACCTTTGAAA	CTGATGAAACAGATAGATTTCTCATGACTTTATTTTGA		2216
Db	2223	AAAGAAATCTTTGAA	ACTGGAACGAAACGGATGAATTTCTCATGAAATTTGTTCTTTGATTA		2282
QY	2217	CAATAAATGTGAAAA	AGGCTGGCTGAAAAACCCAGATATGAAAAATGAAAAGCTGACCA		2276
Db	2283	CAAGAAATGTTGAAAA	AACTAGCTGAAAAACCAAAATATCGAATGAAAAACTCATTTAA		2342
QY	2277	ATTAAGAAATACCAT	ATATGAGCAATATATCTAGGACTGAGGAATCAGCAGAGGAATAT		2336
Db	2343	ATTAAGAAACACGAT	ACTGGAACAAATTCACAAGGTCTGAGGAGTCTCCCGAGGAATAT		2402
QY	2337	CTTTACAAAAACAG	CACAGAGTGCATATGCGCTTTCCCGAGTGAATTACTGAAAAATGAAA		2396
Db	2403	TTTTACAAAAACAG	CACAGACCTACGCACTTTCCAGTGAATCATGAAAAATGCAA		2462
QY	2397	ATTGCTGAAGTAG	AGGTCAAAGCCCAACCATCTGATTTGAGCTGACACAGCAGTGA		2456
Db	2463	GTTTGGGAAGTTG	AGTCAAAGCCCATCACCTGATTTGGCCGGGGCACACAGCAGTGA		2522
QY	2457	CAAAACCATGACA	CAGAAATGAACAAAAAGAGTCAATTAGTAAATTTCCGCACTGAAAAAT		2516
Db	2523	CAAGCCATGACT	CAGACTGAACAAAAAGAGTCAATTAGTAAATTTCCGCACTGGCGAAAT		2582
QY	2517	CAATCTGCTTAT	CGCTACCAAGTGGCAGAGAAGGCTTGATATTAAGAATGTAAACAT		2576
Db	2583	AAATCTGCTTAT	CGCTACGACGCTGGCAGAGAAGGCTTGATATCAAAAGATGCAATAT		2642
QY	2577	TGTTATCCGTTAT	GTCTGTCACCAATGAATAGCCATGGTCCAGGCCGCTGTGAGC		2636
Db	2643	TGTTATCCGTTAT	GTATGCGCTGTTCACGAACGAGATAGCCATGGTCCAGGCCGCGGTGAGC		2702
QY	2637	CAGAGCTGATGA	GACACCTACGTCCTGTTGCTCACAGTGGTTCAAGAGTTATCGAAC		2696
Db	2703	CAGAGCTGATGA	AAAGCACGTAATGTCTTGTCACAGCAGTGGCTCAGGAGTTACCGAACG		2762
QY	2697	TGAGACAGTTA	TGATTTCCGAGAGAAGATGATGTATAAGCTATACATTGTGTTCAAA		2756
Db	2763	GGAGATTGTTA	TGATTTCCGAGAGAAGATGATGTATAAGCTATTAACCGTGTCAAA		2822
QY	2757	TATGAACCA	GAGAGTATGCTCATTAAGATTTTGAATTAAGATGCAAAAGTATTAATGA		2816
Db	2823	CATGAACCA	GAGAGTATGCTCATTAAGATTTTGAATTTGCAAGTATCTGGA		2882
QY	2817	AAAGAAATGA	AAAAACAAGAAATATTGCCAAGCATTAACAAGATAACCCATCACTAAT		2876
Db	2883	AAAGAAATGA	AAAGTCAAAAGAACGATTTGCAAAAGCAATACACGAATCCATCGTTAAT		2942
QY	2877	AACTTTCCTT	TGCAAAAAATCTGCACTGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAAAT		2936
Db	2943	AAACCTTCT	CTGCAAAAAATTTGATGATGTGCTGTGCTGGGAGAAAAATCCATGTAT		3002
QY	2937	TGAGAAAAAT	GATCATGCTCAATATGACCCAGAAATTCAGGAACCTTACATTGTAAAGA		2996
Db	3003	TGAGAAAGAT	CATCATGTCAATATGACACCAAGAAATTCAGGGAACCTTACATTGTAAAGA		3062
QY	2997	AAACAAAGCA	CTGCAAAAGAGTGTGCCGACTATCAAAATAAATGTTGAATCATCTGCA		3056
Db	3063	AAACAAAGCA	CTGCAAAAGAAATTTGCTGATTATCAGACCAATGGAAGATTATCTGCA		3122
QY	3057	ATGTGCCAGG	CTTGGGAACAATGATGTGCACAAAGGCTTAGATTTGCCCTGTCTCA		3116
Db	3123	GTGTGCCAGG	CTTGGGGAACAATGATGTGTGCACAAAGGTTTAGATTTGCCCTGTCTTA		3182
QY	3117	AATAAGGAAT	TTTGTAGTGTTCAAAAATATTTCAACAAAGAAACAATACAAAAAGTG		3176
Db	3183	AATAAGGAAT	TTTGTAGTCAATTTCAAAAAATTACTCACCGAAGAAACAGTACAAAGAGTG		3242

QY	3177	GGTAGAATTACCTATCACATTTC	CCCAATCTTGACTATT	CAGAATGCTGTTATT	TAGTGA	3236
Db	3243	GGTGAATTGCCATACAGATTTC	TGATCTTGACTACT	CAGAATACTGCTGTAT	AGTGA	3302
QY	3237	TGAGGATTAGCACTTGATTGA	AGATTC	TTTAAATACTATCAGTT	AAACA	3296
Db	3303	TGAAGATTAGCACTTGATTCAT	GATTATTTTAAATACTGCCA	ACTCAACATTGAATATG		3362
QY	3297	ATTATGATTAAATGATTCATT	TATGCTACAGAACTGACATA	AGAA	TCAA	3348
Db	3363	ATTTAAATGTTGTTTTCGCTA	CACACTACTGAGCTAACTCT	AGAA	TATATGTA	3414

RESULT 14			
AY075132			
LOCUS	AY075132	3078 bp	mRNA
DEFINITION	Mus musculus HELICARD mRNA, complete cds.	linear	ROD 20-MAY-2002
ACCESSION	AY075132		
VERSION	AY075132.1	GI:18698980	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		

REFERENCE  
AUTHORS  
1 (bases 1 to 3078)  
Kovacsovichs, M., Martinon, F., Michéau, O., Bodmer, J. L., Hofmann, K.

TITLE	JOURNAL	PUBMED
Overexpression of Helicard, a CARD-Containing Helicase Cleaved during Apoptosis, Accelerates DNA Degradation	Curr. Biol. 12 (10), 838-843 (2002)	12015121

**AUTHORS** Kovacsovics, M., Hofmann, K. and Tschopp, J.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (24-JAN-2002) Biochemistry, University of Lausanne, ch  
des Boveresses 155, Epalinges, VD 1066, Switzerland

FEATURES	Location/Qualifiers
Source	1. .3078
	/organism="Mus musculus"
	/mol_type="mRNA"
	/db_xref="taxon:10090"
CDS	1. .3078

/note="hellicase; contains two CARD domains"  
/codon\_start=1  
/product="HELICARD"  
/protein\_id="AAL77205.1"  
/db\_xref="GI:18698981"  
/translation="MSIVCSAEDSFRLILFRRPLKMYIQVEPVLDHLIFLSAETKE  
QLKKINTGCGTSAEALLSTLECGMPLGWTOMFVEALEHSGDNLARVYPTLTDL  
PSPSSETAHDECLHLLTLQPTLVKLLINDVLDTCEFKGLTYEDRNRISAAGNSG  
ESGVELLRIVOKENFSTFLVLRQTDNALFOELTGCGCPENDTLANSHRDGP  
AANECLIPAVDESSLETEAMVNDILPEASCTDSSVTTESDTSIAEGSVSCFDESLGH  
NSNMGDSGTMGSDSDSVIQTRKVSPEPELQLRPYQMEVAQPALDGKNIILICLPTGS  
GKTRVAVYITKDLDDKKQASSEGVYIVLVNKVMAEQLFKEFENPLKKWYRIIGLS  
GDTQLKISFEPEVKS YDVIISTAQILNLSLNLBESGDDGVLDSFLIIDECHHTN  
KEAVYNNIMRRYLKOKLRNNDLKKONKPAIPLPOLGLTASPEGAARKQOSEAEKHIL  
NLCANDLAFITIKTVKENLGQLKHQIKEPCKKFIADDTRENPFKEKLEIMASIQTYC  
OKSPMSDFGTOHYEQWAIOMEKKAADGNRKDRVCAEHLRKXNALQINDTIRIMDAY  
SHLETFTYDEKEKKFAVLNDSKSDDEASSCNDQLKGVDKSLKIDETDEFLMILFED  
NKKMLKLAENPKYENKLIKLRNTILLEQFTRSEBSRGIIFTKRSTQYALISOWIMB  
NAKFAEYGVKAHHLIGAGHSSEYVPMTOEQKEVISKFRTEGINLLIATTVABEGLDI  
KECNIYIRYGLVTNEIAMVOARGARADESTYVLVTSGSGSVTEREIVNDEREKMYK  
AINRVQNMKPEEYAHKILLELOVOSTIEKMKVKRSIAKQYNDNPSLITLLCNCSMLV  
CSGENTHVIKMHVNMTPEFKQYILVRENKALQKKFADYQTNKEIICKCGAWGTMM  
VHKGLDPLCLKIRNFVNFKNNSPKQYKKWVLEPIRFPLDLYSEYCLYSDSD"

## ORIGIN

Query Match	63.5%;	Score 2135.4;	DB 9;	Length 3078;
Best Local Similarity	81.3%;	Pred. No. 0;		
Matches 2504;	Conservative 0;	Mismatches 571;	Indels 6;	Gaps 2;

169 ATGTCGAATGGGTAATCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCC 228

Db	1	ATGTCGATGTCTGTTCTGCAGAGACAGCTTCAGAAATCTCATCTTATTCTTCAGGCC	60
Qy	229	AGGGTGAATAATGTACATTCAGGTGGAGCCTGTGCTGACTAACCTGACCTTCTGCTGCA	288
Db	61	AGCGTAAAAATGTACATTCAAGGTGGAGCCAGTGTCTGACCACTCATCTTCTGTCTGCA	120
Qy	289	GAGGTGAAGGACAGATTCAAGAGACAGTCCGACCTCCGGGAACATGCAGGCGATTGAA	348
Db	121	GAACCAAGACAGATTCTTAAAAAGATCAACACCTGTGTAAACACGCGCGGAGAA	180
Qy	349	CTGCTGCTGAGCACCTTGGAGAAAGGAGTCTGGACACTTGTGTGAACTCGGAATTGCTG	408
Db	181	CTGCTGCTGAGCACCTTGGAGCAATGAGCCTGTGGATGGAACGATGTTGCTG	240
Qy	409	GAGGCCCTCCGAGAACCGGACGCCCTCTGGCCCGCCCTACATGAACCTGAGCTCACG	468
Db	241	GAGGCCCTAGAGCACAGTGGCAATCCCTAGCCCGCGCTATGTCAAAACCACTACT	300
Qy	469	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC	528
Db	301	GATCTGCCCTCTCCTCTCTGAGACTGCCCATGACGAGTGTCTCCACTTGTGACCTC	360
Qy	529	CTTCAGCCCACTCTGTGGACAAGCTTCTAGTTAGAGAGTCTTGATAAGTGCATGAG	588
Db	361	CTCCAGCCCACTTGGTGACAAACTTCTGATTAACGATGCTTGACACTTGTCTGAG	420
Qy	589	GAGGAAGTGTGACAATTGAAGACAGAAACCGAATTGCTCTGCAGAAACCAATGAAAT	648
Db	421	AAAGGACTATTGACAGTCGAGACAGAAATCGAATTTCTGCTGCAGAAACACGCGGAAAT	480
Qy	649	GAATCAGGTGAAGAGAGCTATAAAAGGATTGTGCAGAAAGAAAAGTGTCTCTGCA	708
Db	481	GAGTCAGGTGAAGAGAGCTGTGAGAAAGGATTGTGCAGAAAGAAAAGTGTCTCTGCA	540
Qy	709	TTTCTGAATGTTCTTCTGTCAACAGAAACAATGAACCTTGTCCAAGATTACAGGCTCT	768
Db	541	TTCTGATGTTCTGCGCCAACTGAAATGATGACATTTCCAAGAACTAACAGGTGA	600
Qy	769	GATTGCTCAGAAAGCAATGCAGAGATTGAGATTATCAACAAGTGTATGCTCTCAAGT	828
Db	601	GAGTCCCAAGAAACACACAGACTTGGCTAACTCGTCTACAGAGATTGGCCTGCAGCT	660
Qy	829	GAAAGCAACTCTTTTCAACCAAGTTCAGCCAAATCTGGAAGAGAGGTCTGGGCGATG	888
Db	661	AATGAGTGTCTTCTGCTGTGATGATGATCAAGTCTGAGACAGAGGCTGGAAGTGA	720
Qy	889	GAGAATACTCATCAGAAATCACTTTTGACAGATTCTTCTGTAGTTTCAGAAATCAGACAA	948
Db	721	GACGACATATTACAGAGGCTTCTGTACAGATTGCTGTGACCAAGAAATCAGACAA	780
Qy	949	AGTTTGGCAGAAAGAGTGTCAAGTCTTAGATGAAGTCTTGACATTAACAGCAACATG	1008
Db	781	AGTTTGGCAGAAAGAGTGTCAAGTCTTAGATGAAGTCTTGACATTAACAGCAACATG	840
Qy	1009	GCGAGTGAATCAGGCAACCATGGGAAGTGAATTCAGATGA---GAGAATGTGGCAGCAAGA	1065
Db	841	GCGAGGATTCAGGCAACCATGGGAAGTGAATTCAGATGAAGTGTCAATCCAGACAAAAGA	900
Qy	1066	GCAATCCCGAGAGCCAGAACTCAGCTCAGGCCCTTACCAAAATGGAAGTTGCCAGCCAGCC	1125
Db	901	GTAATCCCGAGAGCCAGAACTCAGCTCAGGCCCTTACCAAAATGGAAGTTGCCAGCCAGCT	960
Qy	1126	TTGGAAGGGAAGATATATCATCATCTGCTCCCTACAGGAGTGAAGAAAACAGAGTGGCT	1185
Db	961	CTAGATGGGAAGATATATATCTGCTCCCTCCACGCGGAGTGGGAAGAAAACAGAGTGGCT	1020
Qy	1186	GTTTACATTTGCCAAGATCATCTTAGACAAGAAAGAAAAGCATCTGAGCCTGGAAGAAATT	1245
Db	1021	GTTTACATTCACCAAGATCATCTTAGACAAGAAAGCAGGCACTGAAATCCGGGAAGGTT	1080
Qy	1246	ATAAGTCTTGTCAATGAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACA	1305

Db	1081	ATCGTTCTTGTCAATTAAGGTAAATGTTAGCAGAACCACTTTCCGAAAAAGAGTTCAACCCA	1140
Qy	1306	TTTTTGAAGAAATGGTATCGTGTTATTGGATTAAAGTGTGATACCCAACTGAAAAATATCA	1365
Db	1141	TATTGGAAGAAATGGTATCGAATTAATTGATTAAAGTGGCATACCCAGCTGAAAAATATCA	1200
Qy	1366	TTTCCAGAAGTTGTCAAGTCTGTGATATTATTATTCAGTACAGCTCAAACTCTTGAAAC	1425
Db	1201	TTTCCAGAAGTTGTCAAAATCTTACGATGTTATTATTCAGCACTGCTCAAACTCTTGAAAC	1260
Qy	1426	TCCCTCTTAAACTTGGAAAAATGGAAGATGTGGTGTCAATTGTTCAGACTTTCCCTC	1485
Db	1261	TCCCTCTTAAATCTGAGAATGGAAGACGATGACGGTGTGACAGTGTCAAGCTTCTCTCTC	1320
Qy	1486	ATTATCATTTGATGATGATCATCACCAACAAGAACAGTGTATATAAATACATCATGAGG	1545
Db	1321	ATTATCATTTGATGATGATGCCATCACCAACAAGAGGAGAGTCTATTAACAATCATGAGA	1380
Qy	1546	CATTATTGATGCAGAAAGTTGAAAAACAATAGACTTCAGAAAGAAAAACAACCAAGTATT	1605
Db	1381	CGATATTGTAAGCAGAAAGCTGAGAAACAATGACCTTCAGAAACAAAACAAGCCATT	1440
Qy	1606	CCCCCTTCTCAGATACTGTGGACTAAACAGCTTCACCTGTGTGGAGGGGCCACGAAGCA	1665
Db	1441	CCCCGCGCAGATACTAGAGCTGACAGCTTCACCTGTGTGGAGCAGCCAAAAAGCAG	1500
Qy	1666	GCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATGATTTACTATTAA	1725
Db	1501	TCTGAGGCTGAAAAACATATTTTAAATATATGTGCCAATCTTGATGCTTTTACATTAA	1560
Qy	1726	ACTGTTAAGAAAACTTGATCAACTGAAAAACCAATAACAGAGCCATGCAAGAAATTT	1785
Db	1561	ACAGTGAAGAGAATCTTGCTCACTCAACACCAATAAAGAAACCATGCAAGAAATTT	1620
Qy	1786	GCCATTGCAGATGCCAACGAGAGAGATCCATTTAAAGAGAACTTTAGAAATAATGACA	1845
Db	1621	GTGATTGCTGATGACACCAAGAGAAAAATCCATTTAAAGAGAACTTTAGAAATAATGCA	1680
Qy	1846	AGGATTCAAACTTATTGTCAAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGAA	1905
Db	1681	AGCATTCAAGCTTACTGCCAAAAAGTCCAAATGTCAGATTTTGGAAACCAACTTAATGAG	1740
Qy	1906	CAATGGGCCATTCAAAATGAAAAAAAAGCTGCAAAAAAGGAAATCGCAAGAAAGTGT	1965
Db	1741	CAGTGGGCCATTCAAAATGAGAAAAAAGCTGTAAGACGAAATCGCAAGATCGCGTC	1800
Qy	1966	TGTGCAAGACATTTGAGGAAGTCAATGAGGCCCTACAAATTAATGACACAAATTCGAATG	2025
Db	1801	TGTGCAAGACATTTGAGGAAGTCAACGAAGCCCTACAAATCAACGACAGATCCGAATG	1860
Qy	2026	ATAGATGCCGTAATCTCATCTTGAACTTTCTATAATGAAGAAAGATAAGAGTTTGCA	2085
Db	1861	ATTGATGCATATAAGCCACCTGGAGACATTTCTACCTGATGAAGAAAGAAAGATTGCA	1920
Qy	2086	GTCATAGAAGATGATAGTATGAGGGGTGTGATGATAGTAATTGTGATGGTGAAGAT	2145
Db	1921	GTCCTC--AATGACAGCGACAGAGATGATGACGAGGCCAGCAGTTGCAATGACCAACTT	1977
Qy	2146	GAGGATGATTTAAAGAAACCTTTGAAACTGAGTGAACAGATAGATTCTCATGACTTTA	2205
Db	1978	AAGGCGATGTAAAGAAATCTTTGAAACTGAGCGAAACGATGAATTTCTCATGAATTTG	2037
Qy	2206	TTTTTTGAAAACAATAAAATGTTGAAAAAGGCTGGCTGAAAAACCAAGATATGAAAAATGAA	2265
Db	2038	TTCTTTGATTAACAAGAAATGTTGAAAAAACTAGCTGAAAAACCAAAATACGAGAAATGAA	2097
Qy	2266	AAGCTGACCAATTAAGAAATACATATGAGCAATATATAGGACTGAGGAATCAGCA	2325
Db	2098	AAACTCATTAATAAAGAAACACAGATACTGGAACAATTCACAAGGTCTGAGAGTCTCTCC	2157
Qy	2326	CGAGGAATTAATCTTTACAAAAACACGACAGAGTGCATATGCCGCTTCCCAAGTGAATTA	2385
Db	2158	CGAGGAATTAATCTTTACAAAAACACGACAGAGCACTTACGCCACTTCCCAAGTGAATCATG	2217



CDS

/db\_xref="MG1:1918836"  
231..3161  
/gene="Ifih1"  
/codon\_start=1  
/product="Ifih1 protein"  
/protein\_id="AAH80200.1"  
/db\_xref="GI:51329784"  
/db\_xref="GeneID:71586"  
/db\_xref="MG1:1918836"  
/translation="MSIVCSAEDSFRLILFRPRLKMYIQVEPVL DHLIFLSAETKE  
QILKIKINTCGNTSAEILLSTLEOGWPLGWTOMFEVALEHSGNPLAARYKPTLIDL  
PSPSETAHDECHLLTLQPTLVKLLINDVLDTCPEKGLTVEDRNRI SAAGSGN  
ESGVRELLRRIJOKEMWESTFLDVLROGNDALFOELTGGCEPDNTESDTSIAEGSV  
SCFDESLGHNSMGRDSTGMSDSDESVIQTKRVSPEPELQRPYQMEVAQPALDGXN  
IICLPTGSGKTRVAVYITKDHLDKKQASESGKIVLVNKKMLAEQLFRKEFNPLYK  
KMYRIIGLSGDTOLKISFEVVKSYDVIISTAOILENSLNLSESGDDGVQLSDFSLI  
IIDECHTNKEAVYNNIMRRLKQKLRNNDLKKQNKPAIPLPQILGLTASPGVGAKK  
QSEAEKHLINICANDAFITIKTKENLGQKHQIKPECKKFIADDTRENPFKEKLE  
IMASIQTYCQKSPMSDFGTQHYEQWALQMEKKAADGNRKDRVCAEHLRKNEALQIN  
DTIRMDAYSHELFYTDEKEKFAVINDSDSDDEASSCNDOLKGDVKKSLKLDETD  
EFLMNLFFDNKKMLKLAENPKYENELIKLRNTILEQFTRESESSRGIIPTKTROST  
YALSQWIMENAKFAEVGVKLAHLIGAGHSEVKPMTOEQKEVLSKERTGEINLIAT  
TVAEGLDIEKCNIVIRYGLVTNEIAMVQARARADESTVYLVTSGSGVTEREIVN  
DREKMMYKAIINRVQNMKPBRYAHKILEQVOSILEKMKVRSIAKQYNDNPSLITL  
LCKNCSMLVCSGENIHVIEKMHVNMTPFKGLYIVRENKALQKFPADYQTNGEIICK  
CGQAWGTMVHKGLDLPCLKIRNFVNVFNKNSPKQYKKMWELPIRFPDLDYSECLY  
SDSD"

ORIGIN

Query Match	59.3%; Score 1995.4; DB 9; Length 3635;
Best Local Similarity	77.3%; Pred. No. 0;
Matches 2544; Conservative	0; Mismatches 591; Indels 154; Gaps 4;
QY	64 GCTTCTCTAAGTGGCAGCGGACAGCGGCACGACATTTCACTGTCCCGCAGACAACAG 123
Db	125 GCAGCTCAGCGCGGACCCCATTCACCTGCACCTGAGATCTCTCCCGCCGCCGCC 184
QY	124 CACCATCTGCTTGGAGAACCTCTCCC-TTCTGTGAGAAAGAAAGATGTGGAATGGTA 182
Db	185 GCCCGCCCGCCTGGAGACTCTCCCATTTCTGTGAGACCTCAGATGTGATGTCTG 244
QY	183 TTCCACAGACGAGAAATTTCCCGTATCTCATCTCGTCTTCAGGGCCAGGGTGAATGTA 242
Db	245 TTCTGACAGAGACAGCTTCAAGSAATCTCATCTTATTCTTCAGGCCCAAGCTGAATAATGTA 304
QY	243 CATCAGGTGAGCCTGTGTGAGACTACCTGACCTTCTGCTGCAGAGGTGAAGGACA 302
Db	305 CATTCAGGTGAGCCAGTGTGGAACCACTCATCTTCTGTCTGAGAAACCAAGAGCA 364
QY	303 GATTGAGAGACAGTGCACCTCCGGAAACATGACGAGCAGTTGAACGTGCTGTGAGCAC 362
Db	365 GATTCTTAAAAAGATCAACACCTGTGTGTAACACGCGCGGACAGAACTGCTGTGAGCAC 424
QY	363 CTTGAGAGAGGAGTGTGGCACTTGTGTGAGCTCGGGAATTGCGTAGGGCCCTCCGAG 422
Db	425 CTTGAGAGAGGAGCAGTGGCTCTGGGATGAGCGCAGATGTTCTGTGAGGGCCCTTAGAGCA 484
QY	423 AACCGGACCCCTCTGGCCCGGCTACATGAACCTGAGCTCACGGAATTGCCCCCTCC 482
Db	485 CAGTGGCAATCCCTAGCCGCGCTATGTCAAAACCAACATCTGATGTGCCCCCTCC 544
QY	483 ATCGTTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAGCCCCACTCT 542
Db	545 TTCTCTGAGACTGCCCATGAGAGTGTCTCCACTGTGACCTCTCCAGCCCACTT 604
QY	543 GGTGACAAAGCTTCTAGTTAGAGCGTCTTGATAGTGCATGAGAGGAACTGTTGAC 602
Db	605 GGTGACAACTTCTGATTACAGATGTCTTGACACTTGTCTCGAAGAGGAACTCTTGAC 664
QY	603 AATTGAGACAGAAACCGGATTGCTGCTGCAGAAAAACAATGGAATGAATCAGGTGAAG 662
Db	665 AGTCGAGACAGAAATCGGATTTCTGCTGCAGGAAACAGCGGGAATGAGTCAAGTGTAAAG 724

QY	663 AGAGCTACTAAAAAGATTTGTGCAGAAAGAAAAGTGTCTCTGCAATTTCTGAATGTTCT 722
Db	725 AGAGCTGCTGAGAGGATTTGTGCAGAAAGAAAAGTGTCTCTGCAATTTCTGATGTTCT 784
QY	723 TCGTCAACAGAGAAACAATGAACCTGTCCAGAGTTAAACAGGCTGTGATGTCTCAGAAAG 782
Db	785 GCGCCAACTGGAATGATGCACTATTCCAGAACTAACAGAGTGAAGGCTGCCAGAGAA 844
QY	783 CAATGACAGATTGAGAAATTTATCACAAGTTGATGCTCCTCAAGTGAAGACCACTTCT 842
Db	845 CAA----- 847
QY	843 TTCAACACAGTTTACGCCAATCTGAGAAAGAGGTCTGGGCAATGAGAAATACTCATC 902
Db	848 ----- 847
QY	903 AGAATCATCTTTTGAGATTCTTCTGTAGTTTCAGAATCAGACACAGATTGGCAGAAAG 962
Db	848 -----CACAGAATCAGACACAGATTGGCAGAAAG 877
QY	963 AAGTGTACGTGCTTAGATGAAGTCTTGACATTAACAGCAACATGGGCAGTGTACAG 1022
Db	878 AAGTGTACGTGCTTGAATGAAGTCTTGACATTAACAGCAACATGGGCAGGATTCAGG 937
QY	1023 CACCATGGGAAGTATTCAGATGAA--GAGAAATGTGGCAGCAAGACATCCCCGAGACC 1079
Db	938 CACCATGGGAAGTATTCAGATGAAAGTGTCAATCCAGCAAAAAAGATATCCCCGAGACC 997
QY	1080 AGAATCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTGGAAGGAGAA 1139
Db	998 AGAATCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCTTAGATGGAGAA 1057
QY	1140 TATCATCATCTGCTCCCTTACAGGAGTGAAGAAAACAGAGTGGCTGTTTACATTGCCAA 1199
Db	1058 TATTATATATCTGCTCCCAAGGAGTGGAGAAACAGAGTGGCTGTTTACATCACCA 1117
QY	1200 GGATCACTTAGACAAAGAAAGAAAGCATCTGAGCCTGAAAAAGTTAAGTCTTGTCAA 1259
Db	1118 AGATCACTTAGACAAAGAAAGACAGCATCTGAATCCGGAGAGTTATGCTTGTCAA 1177
QY	1260 TAAAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAAACCATTTTGAAGAAATG 1319
Db	1178 TAAAGTAAATGTTAGACAAACATTTCCGAAAAGAGTTCAACCCATATTTGAAGAAATG 1237
QY	1320 GTATCGTGTATTTGATTAAGTGTGATACCAACTGAAAATATCATTTCCAGAAAGTTGT 1379
Db	1238 GTATCGAATTAATGATTAAGTGGCGATACCCAGCTGAAAATATCATTTCCAGAAAGTTGT 1297
QY	1380 CAACTCTGTGATATTAATTAAGTACAGTCAATCTTGAAAACTCCCTTAAACTT 1439
Db	1298 CAAATCTTAAGATGTTATTAATCAAGACTGCTCAATCCTTGAAAACTCCCTTAAATCT 1357
QY	1440 GGAATAATGAGAAAGATGCTGTGTTCAATTGTGAGACTTTTCCCTCATTAATCATTGATGA 1499
Db	1358 GGAGATGAGAGACGATGACGCTGTGCAAGCTGTCAAGACTTCTCTCATTAATCATTTGATGA 1417
QY	1500 ATGTCAATCACACCAAGAAAGAGAGTGTATTAATACATCATGAGGCAATTAATTGATGA 1559
Db	1418 GTGCATCACACCAAGAGAGGAGTGTATTAACATCATGAGAGCAATTAATTGAAACA 1477
QY	1560 GAAATTGAAAAACAATAGACTCAAGAAAGAAACAACAGTGAATTCCTTCTCAGAT 1619
Db	1478 GAAGCTGAGAAACAATGACTCAAGAAACAACAACAGCCATTCCTTCCCGCAGAT 1537
QY	1620 ACTGGACTAACAGCTTCACTGTGTGTGAGAGGGCCACGAAGCAAGCCAAAGCTGAAGA 1679
Db	1538 ACTAGACTGACAGCTTCACTGTGTGTGAGAGGCCAAAGAGAGCTGAAGAA 1597
QY	1680 ACACATTTTAAACTATATGTGCCAATCTTGATGCACTTACTATTAACACTGTTAAGAAAA 1739
Db	1598 ACATATTTTAAATATATATGTGCCAATCTTGATGCTTTTACCATTAACAGTGAAGAGAA 1657
QY	1740 CCTTGATCAACTGAAAAACAATAACAGAGAGCATGCAAGAAAGTTTGCCATTGCAAGATGC 1799

Db 1658 TCTTGTCAACTCAAAACACCAATTAAGGAACCATGCAGAAATTTGTGATGCTGATGA 1717  
Qy 1800 AACGAGAGAAGATCCATTAAAGAACTTCTAGAAATTAATGACAAAGATTCAAACTTA 1859  
Db 1718 CACGAGAGAAATCCATTAAAGAGAACTTCTAGAAATTAATGCAAGCATTCAGACTTA 1777  
Qy 1860 TTGTCAATGAGTCCCAATGTGAGATTTTGGAACTCAACCCCTATGACAAATGGGCCATTCA 1919  
Db 1778 CTGCCAAAAAAGTCCCAATGTGAGATTTTGGAACTCCCAACATTAATGAGCAGTGGCCATTCA 1837  
Qy 1920 AATGAAAAAAGCTGCAAAAAAAGGAAATCGCAAGAAACGTGTTGTGCAACAATT 1979  
Db 1838 AATGAGAAAAAAGCTGTAAGAAGGAAATCGCAAGATCGCGTCTGTGCAACAATT 1897  
Qy 1980 GAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTGCAATGATAGATGCGTATAC 2039  
Db 1898 GAGGAAGTACAACGAAGCCCTACAAATCAACGACAGATCCGAATGATGATGATATAG 1957  
Qy 2040 TCATCTTGAACCTTCTTAATGAAGAGAAAGATTAAGATTGCACTGATAGAAAGATGA 2099  
Db 1958 CCACCTGAGACATTTCTACTGATGAGAAAGAAAGATTGCGAGTCT--CAATGA 2014  
Qy 2100 TAGTGATGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2159  
Db 2015 CAGCGACGAGAGTATGATGACGAGGCCAGCAGTTCGATGACCACTTAAGGCGATGTAA 2074  
Qy 2160 GAAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTATTTTGAACAACAA 2219  
Db 2075 GAAATCTTTGAAACTGACGAAACGATGAAATTTCTCATGAAATTTGTTTGAATACAA 2134  
Qy 2220 TAAATGTTGAAAAAGCTGCTGTAACCAAGATATGAATGAAGAGTGAACCAATT 2279  
Db 2135 GAAATGTTGAAAAAAGCTGTAACCAAGATATGAATGAAGAGTGAACCAATT 2194  
Qy 2280 AAGAAATACCATTAATGAGCAATATATAGAGATGAGAAATCAGACGAGCAATATCTT 2339  
Db 2195 AAGAAACACGATACTGAAACAATTCACAAGGTCTGAGAGTCTCCCGAGAAATATTTT 2254  
Qy 2340 TACAAAAACACGACAGAGTGCATATGCGCTTTCCCATGATTAATGAAAAAT 2399  
Db 2255 TACAAAAACACGACAGAGCCTACGACCTTTCCCATGATTAATGAAAAAT 2314  
Qy 2400 TGCTGAAGTAGAGTCAAAAGCCCACTCTGATTTGAGCTGACACAGCAGTGAATTCAA 2459  
Db 2315 TGCGGAAGTTGAGTCAAAAGCGCATCACCTGATTTGCGCGGCGACAGCAGTGAAGTCAA 2374  
Qy 2460 ACCCATGACACAGAAATGAACAAAAAGATCATTAATTTGCACTGCAAAAAATCAA 2519  
Db 2375 GCCCATGACTCAGACTGAACAAAAAGATCATTAATTTGCACTGCGCAAAATAAA 2434  
Qy 2520 TCTGCTTATCGCTACACAGTGGCAGAGAAAGTCTGATATTAAGAATGTAACATTGT 2579  
Db 2435 TCTGCTTATCGCTACAGCGGTGGCAGAGAAAGCTGATATCAAAAGATGCAATATTGT 2494  
Qy 2580 TATCCGTTATGTCCTGTCACCAATGAATAGCCATGCTCCAGGCCCCGTGTCAGCCAG 2639  
Db 2495 TATTCGTTATGCTTGTTCACGAACGAGATAGCCATGCTCCAGGCCCCGTGTCAGCCAG 2554  
Qy 2640 AGCTGATGAGAGCACTTACGTCCTGCTGCTCAGAGTGTTCAGAGATTATCGAAACATGA 2699  
Db 2555 AGCTGATGAAAGCACTTATGTCCTGCTCAGCAGAGTGTTCAGAGATTATCGAAACGGA 2614  
Qy 2700 GACAGTTAATGATTTCCGAGAGAAAGATGATGATTAAGCTATATACATTGTTCAAAATAT 2759  
Db 2615 GATGTTAATGATTTCCGAGAGAAAGATGATGATTAAGCTATTAACCGTGTCAAAACAT 2674  
Qy 2760 GAAACGAGAGAGTATGTCATTAAGATTTTGAATTAACAGATGCAAGTAAATGAAGAAA 2819  
Db 2675 GAAACGAGAGAGTATGTCATTAAGATTTTGAATTAACAGATGCAAGTAAATGAAGAAA 2734  
Qy 2820 GAAATGAAACCAAGAGAAATATTTGCCAAGCATTAACAAGATTAACCATTAATAAC 2879

Db 2735 GAAATGAAAGTCAAAAGAGACATTCGAAAGCAATACACGACATCCATGCTTAATAAC 2794  
Qy 2880 TTTCCTTTCAAAACTGACGTGTGCTAGCCCTGTTCTGGGAGAGATATCCATGTAATGA 2939  
Db 2795 ACTTCTCTGCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2854  
Qy 2940 GAAATGATCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2999  
Db 2855 GAAATGATCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2914  
Qy 3000 CAAAGCACTGCAAAAGAGTGTGCGCACTATCAAAATTAATGATGATGATGATGATGATGAT 3059  
Db 2915 CAAAGCACTGCAAAAGAGTGTGCGCACTATCAAAATTAATGATGATGATGATGATGATGAT 2974  
Qy 3060 TGGCCAGGCTTGGGGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3119  
Db 2975 TGGCCAGGCTTGGGGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3034  
Qy 3120 AAGGAATTTTGTAGTGTGTTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3179  
Db 3035 AAGGAATTTTGTAGTGTGTTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3094  
Qy 3180 AGAATTAACCTATCAGATTTCCCAATCTTGACTATTCAGAAATGCTGTTAATTAATGATGA 3239  
Db 3095 GGAATGCTTATCAGATTTCTGATCTGACTACTCAGAAATGCTGTTAATTAATGATGA 3154  
Qy 3240 GGAATGCTTATCAGATTTCTGATCTGACTACTCAGAAATGCTGTTAATTAATGATGA 3299  
Db 3155 AGAATTAACCTATCAGATTTCTGATCTGACTACTCAGAAATGCTGTTAATTAATGATGA 3214  
Qy 3300 ATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 3348  
Db 3215 TTAATTTGTTTTCGCTACACTGAGCTTAATTAATTAATTAATTAATTAATTAATTAAT 3263

Search completed: March 13, 2006, 23:18:31  
Job time : 16328 secs

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2006, 17:31:31 ; Search time 1894 Seconds  
(without alignments)  
11840.914 Million cell updates/sec

Title: US-09-515-363C-1

Perfect score: 3365

Sequence: 1 gcgcgcgcgcctgagagccc.....aaaatgattgttactctg 3365

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21:\*  
1: \_geneseqn1980s:\*  
2: \_geneseqn1990s:\*  
3: \_geneseqn2000s:\*  
4: \_geneseqn2001as:\*  
5: \_geneseqn2001bs:\*  
6: \_geneseqn2002as:\*  
7: \_geneseqn2002bs:\*  
8: \_geneseqn2003as:\*  
9: \_geneseqn2003bs:\*  
10: \_geneseqn2003cs:\*  
11: \_geneseqn2003ds:\*  
12: \_geneseqn2004as:\*  
13: \_geneseqn2004bs:\*  
14: \_geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3365	100.0	3365	4	AAD17203	Aad17203 Human mel
2	3365	100.0	3380	12	ADJ74904	Adj74904 Marker ge
3	3365	100.0	3380	13	ADR14404	Adr14404 Human NF-
4	3365	100.0	3380	14	AEA23691	Aea23691 Human PRO
5	3365	100.0	3380	14	AEA36113	Aea36113 Human nuc
6	3361.8	99.9	3434	14	ADY17563	Ady17563 DNA encod
7	3361.8	99.9	3434	14	ADY20376	Ady20376 DNA encod
8	3360.2	99.9	3928	11	ACN91892	Acn91892 Breast ca
9	3353	99.6	3379	12	ADN04879	Adn04879 Antipsori
10	3339.2	99.2	3668	12	ADQ22237	Adq22237 Human sof
11	3338.8	99.2	3372	6	ABA04908	Aba04908 Human RNA
12	3316.2	98.5	3446	10	ADC30823	Adc30823 Human nov
13	3104	92.2	3131	4	AAD11170	Aad11170 Human mel
14	2202.4	65.5	3771	12	ADJ75813	Adj75813 Marker ge
15	1878	55.8	1967	4	AAS40960	Aas40960 cDNA enco
16	1766	52.5	1776	10	ADA53260	Ada53260 Human cod
17	1460.2	43.4	1557	4	AA159285	Aa159285 Human pol
18	1435.6	42.7	1443	6	ABA04916	Aba04916 Human pol
19	1316.6	39.1	1382	4	AA161071	Aa161071 Human pol

20	1316.6	39.1	1382	10	ADC32572	Adc32572 Human nov
21	1278.8	38.0	1392	11	ADI30722	Adi30722 Human cdn
22	1278.8	38.0	1392	13	ADS82789	Ads82789 Human lym
23	1254.6	37.3	1284	6	ABA04913	Aba04913 Human RNA
24	1130.8	33.6	1319	4	AAS41517	Aas41517 cDNA enco
25	803	23.9	1258	4	AAS01149	Aas01149 Interfero
26	803	23.9	1270	4	AAS01151	Aas01151 Interfero
27	738.8	22.0	956	5	AAS91688	Aas91688 DNA encod
28	663	19.7	3692	4	AAS01150	Aas01150 Interfero
29	663	19.7	3704	4	AAS01152	Aas01152 Interfero
30	577.2	17.2	609	6	ABQ54189	Abq54189 Human ova
31	577.2	17.2	609	6	ABQ54189	Abq54189 Human ova
32	556	16.5	6406	4	AAD17207	Aad17207 Human mel
33	554	16.5	1036	4	AAD17204	Aad17204 Human mel
34	507.2	15.1	595	4	ABA61000	Ab61000 Human foe
35	507.2	15.1	595	4	AA140897	Aa140897 Probe #95
36	507.2	15.1	595	4	AAK35181	Aak35181 Human bon
37	507.2	15.1	595	4	AAK09292	Aak09292 Human bra
38	455.2	13.5	499	4	ABA73600	Ab73600 Human foe
39	455.2	13.5	499	4	AA154039	Aa154039 Probe #22
40	455.2	13.5	499	4	AAK48209	Aak48209 Human bon
41	455.2	13.5	499	4	AAK22046	Aak22046 Human bra
42	453	13.5	453	12	ACH90906	Ach90906 Human gen
43	419.4	12.5	457	9	ACH25741	Ach25741 Human adu
44	376.4	11.2	458	3	AAA64601	Aaa64601 Nucleotid
45	371.2	11.0	392	5	AAF66485	Aaf66485 Novel hum

ALIGNMENTS

RESULT 1	
ID	AAD17203 standard, cDNA; 3365 BP.
XX	AC AAD17203;
XX	DT 29-NOV-2001 (first entry)
XX	DE Human melanoma differentiation associated (Mda)-5 cDNA.
KW	Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;
KW	RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
KW	neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
KW	breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
KW	central nervous system; cystostatic; apoptosis; ss.
OS	Homo sapiens.
XX	XX
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	/*tag= a
FT	/product= "Human melanoma differentiation associated
FT	(Mda)-5 protein"
FT	3225..3328
FT	/*tag= b
FT	/note= "ATTTA motif"
FT	3284..3387
FT	/*tag= c
FT	/note= "ATTTA motif"
FT	3343..3348
FT	/*tag= d
PN	WO200164707-A1.
XX	XX
PD	07-SEP-2001.
XX	XX
PF	28-FEB-2001; 2001WO-US006960.
XX	XX
PR	29-FEB-2000; 2000US-00515363.
XX	XX
PA	(UYCO ) UNIV COLUMBIA NEW YORK.
XX	XX

PI Fisher PB, Kang D, Gopalkrishnan RV;  
XX  
DR WPI, 2001-565494/63.  
DR P-PSDB; AAEI0155.  
XX  
PT Nucleic acid sequences encoding a Melanoma Differentiation Associated  
PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral  
PT activity.  
XX  
PS Claim 1; Page 16-18; 152pp; English.  
XX  
CC The present invention relates to an isolated nucleic acid encoding a  
CC melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5  
CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.  
CC Mda-5 is a novel interferon (IFN) inducible gene with structural  
CC similarities to RNA helicases and CARD motif containing proteins. Mda-5  
CC is induced during terminal differentiation in human melanoma cells  
CC treated with the combination of recombinant fibroblast IFN and the  
CC antileukaemic compound mezerein (MEZ). Mda-5 is useful for identifying  
CC compounds that may induce its expression. Mda-5 is useful for treating  
CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma  
CC multiforme, cervical cancer, breast cancer, colon cancer, prostate  
CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a  
CC cancer of the central nervous system and apoptosis. The Mda-5 promoter  
CC exhibits melanocyte tissue specificity and minimises systemic toxicity.  
CC The present sequence is human Mda-5 cDNA  
CC  
XX  
SQ Sequence 3365 BP; 1138 A; 644 C; 753 G; 830 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 3365; DB 4; Length 3365;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GCGGCGCCGGCCTGAGAGCCCTGTGGACAACCTCGTCATTGTCAAGGCACAGACCGGTAGAC 60  
DB 1 GCGGCGCCGGCCTGAGAGCCCTGTGGACAACCTCGTCATTGTCAAGGCACAGACCGGTAGAC 60  
QY 61 CTTGCTTCTCTAAGTGGGACGCGACAGCGGCACGACATTTCACCTGTCCGACAGACAA 120  
DB 61 CTTGCTTCTCTAAGTGGGACGCGACAGCGGCACGACATTTCACCTGTCCGACAGACAA 120  
QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTGAGAAAGAAAGATGTGAATGG 180  
DB 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTGAGAAAGAAAGATGTGAATGG 180  
QY 181 TATTTCCACAGACGAGAATTCGGCTATCTCATCTGCTTCAAGGGCCAGGGTGAATG 240  
DB 181 TATTTCCACAGACGAGAATTCGGCTATCTCATCTGCTTCAAGGGCCAGGGTGAATG 240  
QY 241 TACATCCAGGTGAGGCTGTGCTGACTACCTGACCTTCTGCTGCAAGGTGAAGAG 300  
DB 241 TACATCCAGGTGAGGCTGTGCTGACTACCTGACCTTCTGCTGCAAGGTGAAGAG 300  
QY 301 CAGATTCAGAGACAGTCGCCACCTCCGGGAACATGCAAGCAGTTGAATGCTGTAGC 360  
DB 301 CAGATTCAGAGACAGTCGCCACCTCCGGGAACATGCAAGCAGTTGAATGCTGTAGC 360  
QY 361 ACCTTGAGAGAGGAGTGTGGCACTTGTTGACTCGGAATTCGTGAGGCCCTCCGG 420  
DB 361 ACCTTGAGAGAGGAGTGTGGCACTTGTTGACTCGGAATTCGTGAGGCCCTCCGG 420  
QY 421 AGAACCGGACGCCCTGTGGCGCCCGCTACATGAACCTGAGCTCACGGACTTGCCCTCT 480  
DB 421 AGAACCGGACGCCCTGTGGCGCCCGCTACATGAACCTGAGCTCACGGACTTGCCCTCT 480  
QY 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGTGAACCTCTTCAGCCACT 540  
DB 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGTGAACCTCTTCAGCCACT 540  
QY 541 CTGGTGACAAGCTTCTAGTTAGAGACGCTTGTGAATAAGTGCATGAGAGGAACTGTTG 600  
DB 541 CTGGTGACAAGCTTCTAGTTAGAGACGCTTGTGAATAAGTGCATGAGAGGAACTGTTG 600

QY 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAAACAATGGAATGAATCAGGTGA 660  
DB 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAAACAATGGAATGAATCAGGTGA 660  
QY 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAACGTGTTCTCTGCATTTCTGAATGTT 720  
DB 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAACGTGTTCTCTGCATTTCTGAATGTT 720  
QY 721 CTTGCTCAACACAGAAACAATGAACCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA 780  
DB 721 CTTGCTCAACACAGAAACAATGAACCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA 780  
QY 781 AGCAATGCAGAGATTGAGAAATTATCACAAGTTGATGTCCTCAAGTGAAGACAACTT 840  
DB 781 AGCAATGCAGAGATTGAGAAATTATCACAAGTTGATGTCCTCAAGTGAAGACAACTT 840  
QY 841 CTTTCAACCACAGTTCAGCCAAATCTGGAAGAGAGGTCTGGGCGATGGAATTAATCTCA 900  
DB 841 CTTTCAACCACAGTTCAGCCAAATCTGGAAGAGAGGTCTGGGCGATGGAATTAATCTCA 900  
QY 901 TCAGAAATCATCTTTGCAGATTCTTCTGTAGTTTCAAGAAATCAGACACAAGTTGGCAGAA 960  
DB 901 TCAGAAATCATCTTTGCAGATTCTTCTGTAGTTTCAAGAAATCAGACACAAGTTGGCAGAA 960  
QY 961 GGAAGTGCAGCTGCTTAGATGAAGTCTTGACATACAGACAATGGGCAGTGAATTC 1020  
DB 961 GGAAGTGCAGCTGCTTAGATGAAGTCTTGACATACAGACAATGGGCAGTGAATTC 1020  
QY 1021 GGCAACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGACATCCCCGAGCCA 1080  
DB 1021 GGCAACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGACATCCCCGAGCCA 1080  
QY 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTGGCCAGCCAGCCTTGGAAAGGAAGAAAT 1140  
DB 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTGGCCAGCCAGCCTTGGAAAGGAAGAAAT 1140  
QY 1141 ATCATCATCTGCTCCCTTACAGGAGTGGAAAAACAGAGTGGCTTTTACATTGCCAAG 1200  
DB 1141 ATCATCATCTGCTCCCTTACAGGAGTGGAAAAACAGAGTGGCTTTTACATTGCCAAG 1200  
QY 1201 GATCACTTAGACACAGAAAAAAGCATGTGAGCCTGGAAGAAGTTATAGTTCTGTCAAT 1260  
DB 1201 GATCACTTAGACACAGAAAAAAGCATGTGAGCCTGGAAGAAGTTATAGTTCTGTCAAT 1260  
QY 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGGAAGAGTTCCAACCAATTTTGAAGAAATGG 1320  
DB 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGGAAGAGTTCCAACCAATTTTGAAGAAATGG 1320  
QY 1321 TATCGTGTATTGGAATTAAGTGTGATACCCTGAAATATATATTCAGAAAGTTGTC 1380  
DB 1321 TATCGTGTATTGGAATTAAGTGTGATACCCTGAAATATATATTCAGAAAGTTGTC 1380  
QY 1381 AAGTCTGTGATATATATATCAGTACAGCTCAAACTCTTGAAGAACTCCCTTAAACTTG 1440  
DB 1381 AAGTCTGTGATATATATATCAGTACAGCTCAAACTCTTGAAGAACTCCCTTAAACTTG 1440  
QY 1441 GAAATGAGAGAGATGCTGTGTTCAATGTTCAGACTTTTCCCTCATTTATCATTGATGAA 1500  
DB 1441 GAAATGAGAGAGATGCTGTGTTCAATGTTCAGACTTTTCCCTCATTTATCATTGATGAA 1500  
QY 1501 TGTCAATCACACCAAGAAAGAGCAGTGTATTAATACATCATGAGGCATTAATTGATGCAG 1560  
DB 1501 TGTCAATCACACCAAGAAAGAGCAGTGTATTAATACATCATGAGGCATTAATTGATGCAG 1560  
QY 1561 AAGTGAAGAAACAATAGACTCAAGAAAGAAACAACCAAGTGAATTCCTCTCAGATA 1620  
DB 1561 AAGTGAAGAAACAATAGACTCAAGAAAGAAACAACCAAGTGAATTCCTCTCAGATA 1620  
QY 1621 CTGGGACTAACAGCTTCACTGTGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
DB 1621 CTGGGACTAACAGCTTCACTGTGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
QY 1681 CACATTTTAAACATATGTGCCAATCTTGATGATTTACTATTAAGAACTGTAAAGAAAC 1740

Db 1681 CACATTTAAACTATGTCCTCAATCTTGATGCAATTACTATTAAACTGTTAAAGAAAAC 1740  
Qy 1741 CTGTATCACTGAAAAAACAATACAGAGCCATGCAAGAAGTTGCCATTCAGATGCA 1800  
Db 1741 CTGTATCACTGAAAAAACAATACAGAGCCATGCAAGAAGTTGCCATTCAGATGCA 1800  
Qy 1801 ACCAGAGAAGATCCATTTAAAGAGAACTTCTAGAATAATGACAAGATTCAACTTAT 1860  
Db 1801 ACCAGAGAAGATCCATTTAAAGAGAACTTCTAGAATAATGACAAGATTCAACTTAT 1860  
Qy 1861 TGTCAATAGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGAAACAATGGCCATTCA 1920  
Db 1861 TGTCAATAGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGAAACAATGGCCATTCA 1920  
Qy 1921 ATGGAAGAAAAAGCTGCAAAAAAGGAATCGCAAGAACGTTGTTGTCAGAACATTGG 1980  
Db 1921 ATGGAAGAAAAAGCTGCAAAAAAGGAATCGCAAGAACGTTGTTGTCAGAACATTGG 1980  
Qy 1981 AGGAAGTACAATGAGGCCCTACAATTAATGACACAATTGCAATGATGATGCGTATACT 2040  
Db 1981 AGGAAGTACAATGAGGCCCTACAATTAATGACACAATTGCAATGATGATGCGTATACT 2040  
Qy 2041 CATCTTGAACCTTCTATTAATGAGAAGAAAGATTAAGAAGTTGCAGTCAATAGAATGAT 2100  
Db 2041 CATCTTGAACCTTCTATTAATGAGAAGAAAGATTAAGAAGTTGCAGTCAATAGAATGAT 2100  
Qy 2101 AGTGATGAGGGTGGTGATGATGAGTATTTGATGGTGATGAAGATGAGATGATTTAAG 2160  
Db 2101 AGTGATGAGGGTGGTGATGATGAGTATTTGATGGTGATGAAGATGAGATGATTTAAG 2160  
Qy 2161 AAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGAACAAT 2220  
Db 2161 AAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGAACAAT 2220  
Qy 2221 AAAATGTTGAAAAAGCTGCTGAAAAACCCAGAAATATGAAAAAGCTGACCAATTA 2280  
Db 2221 AAAATGTTGAAAAAGCTGCTGAAAAACCCAGAAATATGAAAAAGCTGACCAATTA 2280  
Qy 2281 AGAATAACCATATATGAGCAATATATAGAGCTGAGGAATCAGACGAGAAATATCTTT 2340  
Db 2281 AGAATAACCATATATGAGCAATATATAGAGCTGAGGAATCAGACGAGAAATATCTTT 2340  
Qy 2341 ACAAAAAACAGACAGAGTGATATGCGCTTCCAGTGAGTTACTGAAAATGAAAAATTT 2400  
Db 2341 ACAAAAAACAGACAGAGTGATATGCGCTTCCAGTGAGTTACTGAAAATGAAAAATTT 2400  
Qy 2401 GCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTTGAGCTGACACAGCAGTGAAGTTCAA 2460  
Db 2401 GCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTTGAGCTGACACAGCAGTGAAGTTCAA 2460  
Qy 2461 CCCATGACACAGAAATGAACAAAAGAGTCAATTAATAATTTCCAGCTGAAAAATCAAT 2520  
Db 2461 CCCATGACACAGAAATGAACAAAAGAGTCAATTAATAATTTCCAGCTGAAAAATCAAT 2520  
Qy 2521 CTGCTTATCGCTACCAACAGTGGCAGAAAGGCTTGCAATTAAGAAGTGAACATTTGT 2580  
Db 2521 CTGCTTATCGCTACCAACAGTGGCAGAAAGGCTTGCAATTAAGAAGTGAACATTTGT 2580  
Qy 2581 ATCCGTTATGCTCTGTCACCAATGAATAATAGCCATGGTCCAGGCCGCTGGAGCCAGA 2640  
Db 2581 ATCCGTTATGCTCTGTCACCAATGAATAATAGCCATGGTCCAGGCCGCTGGAGCCAGA 2640  
Qy 2641 GCTGATGAGAGCACCTACGCTCTGCTGCTCAGAGTGGTTCAGAGTTATCGAACATGAG 2700  
Db 2641 GCTGATGAGAGCACCTACGCTCTGCTGCTCAGAGTGGTTCAGAGTTATCGAACATGAG 2700  
Qy 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTTGTTCAAAATATG 2760  
Db 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTTGTTCAAAATATG 2760  
Qy 2761 AAACGAGAGAGTATGCTCATAGAATTTTGAATTAAGATGCAAGTATATGGAAG 2820  
Db 2761 AAACGAGAGAGTATGCTCATAGAATTTTGAATTAAGATGCAAGTATATGGAAG 2820

Db 2761 AAACGAGAGAGTATGCTCATAGAATTTTGAATTAAGATGCAAGTATATGGAAG 2820  
Qy 2821 AAAATGAAAAACAGAGAATAATATTTGCCAGCATTTACAAGAATAACCATCAATAACT 2880  
Db 2821 AAAATGAAAAACAGAGAATAATATTTGCCAGCATTTACAAGAATAACCATCAATAACT 2880  
Qy 2881 TTCCTTTGCAAAAACTGCAAGTGTGCTAGCTGTCTCTGGGGAAGATATCCATGTAATTGAG 2940  
Db 2881 TTCCTTTGCAAAAACTGCAAGTGTGCTAGCTGTCTCTGGGGAAGATATCCATGTAATTGAG 2940  
Qy 2941 AAAATGCATCACGTCATATATGACCCCGAGATTCAAGGAATTTACATTTGTAAGAAAAAC 3000  
Db 2941 AAAATGCATCACGTCATATATGACCCCGAGATTCAAGGAATTTACATTTGTAAGAAAAAC 3000  
Qy 3001 AAAGCACTGCAAAAGAGTGTGCGGACTATCAATAAATGTTGAATCTGCAAAATGT 3060  
Db 3001 AAAGCACTGCAAAAGAGTGTGCGGACTATCAATAAATGTTGAATCTGCAAAATGT 3060  
Qy 3061 GGCCAGGCTTGGGGAACAATGATGCTGCAAAAGGCTTAGATTGCTGCTCAAAATA 3120  
Db 3061 GGCCAGGCTTGGGGAACAATGATGCTGCAAAAGGCTTAGATTGCTGCTCAAAATA 3120  
Qy 3121 AGGAATTTGTAGTGTGTTTCAAAAATATTTCAACAAGAAACAATCAAAAAAGTGCGTA 3180  
Db 3121 AGGAATTTGTAGTGTGTTTCAAAAATATTTCAACAAGAAACAATCAAAAAAGTGCGTA 3180  
Qy 3181 GAATTAACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTTAGTATGAG 3240  
Db 3181 GAATTAACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTTAGTATGAG 3240  
Qy 3241 GATTAGCACTGATTGAAGATTTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300  
Db 3241 GATTAGCACTGATTGAAGATTTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300  
Qy 3301 TGATTAATGTATTCATTAATGCTACAGAACTGACATAAGAATCAATAAATGATGTTTA 3360  
Db 3301 TGATTAATGTATTCATTAATGCTACAGAACTGACATAAGAATCAATAAATGATGTTTA 3360  
Qy 3361 CTCTG 3365  
Db 3361 CTCTG 3365

RESULT 2  
ADJ74904  
ID ADJ74904 standard; DNA; 3380 BP.  
XX AC ADJ74904;  
XX AC  
DT 20-MAY-2004 (first entry)  
XX DE  
DE Marker gene SEQ ID NO:156.  
XX KW bronchial asthma; chronic obstructive pulmonary disease;  
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;  
KW gene therapy; marker gene; gene; ds.  
OS Homo sapiens.  
XX PN EPI394274-A2.  
XX PD  
PD 03-MAR-2004.  
XX PF 04-AUG-2003; 2003BP-00254857.  
XX PR 06-AUG-2002; 2002JP-00229312.  
PR 20-MAR-2003; 2003JP-00077212.  
XX PA (GENO-) GENOX RES INC.  
XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;  
XX DR WPI; 2004-193155/19.

XX Testing for bronchial asthma or chronic obstructive pulmonary disease by  
PT comparing the expression level of a marker gene in a biological sample  
PT from a subject with the expression level of the gene in a sample from a  
PT healthy subject.  
XX  
PS Claim 1; SEQ ID NO 156; 241bp; English.  
XX  
CC The present invention describes a method of testing for bronchial asthma  
CC or chronic obstructive pulmonary disease. The method comprises  
CC determining the expression level of a marker gene in a biological sample  
CC from a subject, comparing the expression level determined with the  
CC expression level of the marker gene in a biological sample from a healthy  
CC subject, and judging whether the subject has bronchial asthma or chronic  
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of  
CC genes (S1) whose expression levels increase when respiratory epithelial  
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)  
CC whose expression levels decrease when respiratory epithelial cells are  
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for  
CC testing for bronchial asthma or chronic obstructive pulmonary disease;  
CC (2) a kit for screening for a candidate compound for a therapeutic agent  
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)  
CC an animal model for bronchial asthma or chronic obstructive pulmonary  
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a  
CC method for producing an animal model for bronchial asthma or chronic  
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial  
CC asthma or chronic obstructive pulmonary disease, comprising the compound,  
CC a marker gene or an antisense nucleic acid corresponding to a portion of  
CC the marker gene, a ribozyme, a polynucleotide that suppresses the  
CC expression of the gene through an RNAi effect or an antibody recognising  
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for  
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a  
CC probe has been immobilised to assay a marker gene. (I) has respiratory  
CC and antiasthmatic activities, and can be used in gene therapy. The method  
CC is useful for testing for or screening for a therapeutic agent for  
CC bronchial asthma or chronic obstructive pulmonary disease. The present  
CC sequence is used in the exemplification of the present invention.  
CC  
XX  
SQ Sequence 3380 BP; 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 3365; DB 12; Length 3380;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 421 AGAACCGGACGCTCTGGCCGCCCTACATGAACCTGAGCTCACGGACTTGCCCTCT 480  
Qy 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCTTGAGCCCACT 540  
Db 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCTTGAGCCCACT 540  
Qy 541 CTGTGACAAAGCTTCTAGTAGACGCTTGATGAATGATGAGAGGAAGTCTTG 600  
Db 541 CTGTGACAAAGCTTCTAGTAGACGCTTGATGAATGATGAGAGGAAGTCTTG 600  
Qy 601 ACAATTGAAGACAGAAACCGGATTGCTGCTGAGAAAACAATGSAATGAAATCAGGTGA 660  
Db 601 ACAATTGAAGACAGAAACCGGATTGCTGCTGAGAAAACAATGSAATGAAATCAGGTGA 660  
Qy 661 AGAGAGCTACTAAAAAGGATTGTGACAGAAAGAAACTGTTCTCTGCATTTCTGAATGT 720  
Db 661 AGAGAGCTACTAAAAAGGATTGTGACAGAAAGAAACTGTTCTCTGCATTTCTGAATGT 720  
Qy 721 CTTCGTCAAAACAGGAAACAATGAACCTTGCCAAAGATTAAACAGGCTTGATTGCTCAGAA 780  
Db 721 CTTCGTCAAAACAGGAAACAATGAACCTTGCCAAAGATTAAACAGGCTTGATTGCTCAGAA 780  
Qy 781 AGCAATGACAGATTGAGAAATTTATCAAAAGTTGATGTCCTCAAGTGAAGAGCAACTT 840  
Db 781 AGCAATGACAGATTGAGAAATTTATCAAAAGTTGATGTCCTCAAGTGAAGAGCAACTT 840  
Qy 841 CTTTCAACCAAGTTCAGCCAAATCTGAGAAAGAGTCTGGGCAATGAGAAATTAATCTCA 900  
Db 841 CTTTCAACCAAGTTCAGCCAAATCTGAGAAAGAGTCTGGGCAATGAGAAATTAATCTCA 900  
Qy 901 TCAGAAATCATCTTTTGACAGATTCTTCTGTAATTTCAAGAAATCAACAAGTTTGACAGAA 960  
Db 901 TCAGAAATCATCTTTTGACAGATTCTTCTGTAATTTCAAGAAATCAACAAGTTTGACAGAA 960  
Qy 961 GGAAGTGTACAGCTGCTTAGATGAAGAGTCTTGACATAACAGCAACATGGGCAATGATCTCA 1020  
Db 961 GGAAGTGTACAGCTGCTTAGATGAAGAGTCTTGACATAACAGCAACATGGGCAATGATCTCA 1020  
Qy 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGAGCCA 1080  
Db 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGAGCCA 1080  
Qy 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCTTGGAAGGGAAGAAAT 1140  
Db 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCTTGGAAGGGAAGAAAT 1140  
Qy 1141 ATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGGCTTTACATTGCCAAG 1200  
Db 1141 ATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGGCTTTACATTGCCAAG 1200  
Qy 1201 GATCACTTAGACAAGAAAAAGCATCTGAGCCTGGAAGAAATTAAGTTCTGTCAAT 1260  
Db 1201 GATCACTTAGACAAGAAAAAGCATCTGAGCCTGGAAGAAATTAAGTTCTGTCAAT 1260  
Qy 1261 AAGTACTGTAGTTGAACAGCTCTTCGCAAGGAGTTCCAACTTTTGAAGAAATGG 1320  
Db 1261 AAGTACTGTAGTTGAACAGCTCTTCGCAAGGAGTTCCAACTTTTGAAGAAATGG 1320  
Qy 1321 TATCGTGTATGATTAAGTGTGATACCCAACTGAAAAATATCATTTCCAGAAAGTTGTC 1380  
Db 1321 TATCGTGTATGATTAAGTGTGATACCCAACTGAAAAATATCATTTCCAGAAAGTTGTC 1380  
Qy 1381 AAGTCTGTGATATTATATATCAGTACAGCTCAATCCTTGAAAACTCCCTTAACCTTG 1440  
Db 1381 AAGTCTGTGATATTATATATCAGTACAGCTCAATCCTTGAAAACTCCCTTAACCTTG 1440  
Qy 1441 GAAAAATGAGAAAGATGCTGCTGTTCAATTGTCAGACTTTTCCCTCATTTATCATTTGATGA 1500  
Db 1441 GAAAAATGAGAAAGATGCTGCTGTTCAATTGTCAGACTTTTCCCTCATTTATCATTTGATGA 1500  
Qy 1501 TGTCAATCACACCAACAAAGAAAGCAGTGTATTAATACATCATGAGCATTAATTGATGACAG 1560  
Db 1501 TGTCAATCACACCAACAAAGAAAGCAGTGTATTAATACATCATGAGCATTAATTGATGACAG 1560

Db	1501	TGTCATCACACCAACAAAGAGCAGTGTATTAATTAACATCATGAGGCATTATTGATGCAG	1560
Qy	1561	AAGTGAACCAATAAGCTCAAGAAAGAAAACAAACAGTGAATCCCCTTCAGATA	1620
Db	1561	AAGTTGAAAAACAATAGACTCAAGAAAGAAAACAAACAGTGAATCCCCTTCAGATA	1620
Qy	1621	CTGGAGCTAACAGCTTCACTGTGTGTGGAGGGGCCAAGCAAGCCAAAGCTGAAGAA	1680
Db	1621	CTGGAGCTAACAGCTTCACTGTGTGTGGAGGGGGCCACGAAGCAAGCCAAAGCTGAAGAA	1680
Qy	1681	CACATTTTAAACTATGTGCCAATCTTGATGCAATTAATAAACTGTTAAAGAAAC	1740
Db	1681	CACATTTTAAACTATGTGCCAATCTTGATGCAATTAATAAACTGTTAAAGAAAC	1740
Qy	1741	CTTGATCAACTGAAAAACCAATAACAGAGCCATGCAAGAGTTGCCATTGCAATGCA	1800
Db	1741	CTTGATCAACTGAAAAACCAATAACAGAGCCATGCAAGAGTTGCCATTGCAATGCA	1800
Qy	1801	ACCAGAGAGATCCATTTAAAGAAACTTCTAGAAATAAGCAAGAGATTCAACTTAT	1860
Db	1801	ACCAGAGAGATCCATTTAAAGAGAACTTCTAGAAATAAGCAAGAGATTCAACTTAT	1860
Qy	1861	TGTCAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCTATGAACAATGGGCCATTCAA	1920
Db	1861	TGTCAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCTATGAACAATGGGCCATTCAA	1920
Qy	1921	ATGGAACCAAGCTGCAACCAAGAAATCGCAAGAAAGCTGTTGTGCAAGCAATTTG	1980
Db	1921	ATGGAACCAAGCTGCAACCAAGAAATCGCAAGAAAGCTGTTGTGCAAGCAATTTG	1980
Qy	1981	AGGAAGTACATGAGGCCCTACAATTAATGACACAAATTCGAATGATAGTGCCTACT	2040
Db	1981	AGGAAGTACATGAGGCCCTACAATTAATGACACAAATTCGAATGATAGTGCCTACT	2040
Qy	2041	CATCTGAAACTTTCTATTAATGAAGAGAAAGATAAGAGTTTGCACTATAGAGATGAT	2100
Db	2041	CATCTGAAACTTTCTATTAATGAAGAGAAAGATAAGAGTTTGCACTATAGAGATGAT	2100
Qy	2101	AGTGATGAGGGTGTGATGATGAGTATTGTGATGGTGAAGAGTGAGATGATTTAAAG	2160
Db	2101	AGTGATGAGGGTGTGATGATGAGTATTGTGATGGTGAAGAGTGAGATGATTTAAAG	2160
Qy	2161	AAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTATTTTGAACCAAT	2220
Db	2161	AAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTATTTTGAACCAAT	2220
Qy	2221	AAATGTTGAAAGGCTGGCTGAAAAACCCAGAAATATGAATAAGCTGACCAATTA	2280
Db	2221	AAATGTTGAAAGGCTGGCTGAAAAACCCAGAAATATGAATAAGCTGACCAATTA	2280
Qy	2281	AGAAATACCATTAATGAGCAATATATACTAGAGCTGAGGAATCAGCAGAGAAATATCTTT	2340
Db	2281	AGAAATACCATTAATGAGCAATATATACTAGAGCTGAGGAATCAGCAGAGAAATATCTTT	2340
Qy	2341	ACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGCATTACTGAAATGAAAAATTT	2400
Db	2341	ACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGCATTACTGAAATGAAAAATTT	2400
Qy	2401	GCTGAAGTAGAGTCAAAAGCCCAACATCTGATTGAGCTGGAACAGCAGTGAATTCAAA	2460
Db	2401	GCTGAAGTAGAGTCAAAAGCCCAACATCTGATTGAGCTGGAACAGCAGTGAATTCAAA	2460
Qy	2461	CCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAAATTTGCACTGAAAAATCAAT	2520
Db	2461	CCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAAATTTGCACTGAAAAATCAAT	2520
Qy	2521	CTGCTTATCGCTACCAAGTGGCAGAAGAGTCTGATATTAAGAATGTAACTTTGTT	2580
Db	2521	CTGCTTATCGCTACCAAGTGGCAGAAGAGTCTGATATTAAGAATGTAACTTTGTT	2580
Qy	2581	ATCCGTTATGTTCTGTCACCAATGAATAGCCATGGTTCAGGCCCGTGTGAGCCAGA	2640
Db	2581	ATCCGTTATGTTCTGTCACCAATGAATAGCCATGGTTCAGGCCCGTGTGTGAGCCAGA	2640

OY		2641	GCTGATGAGACACTACGTCCTGGTTGTCCACAGTGGTTCAAGAGTTATCGAACATGAG	2700			
Db		2641	GCTGATGAGACACTACGTCCTGGTTGTCCACAGTGGTTCAAGAGTTATCGAACATGAG	2700			
OY		2701	ACAGTTAATGATTTCGAGAGAAGATGATGATAAACTATACATTGTGTCAAATAATG	2760			
Db		2701	ACAGTTAATGATTTCGAGAGAAGATGATGATAAACTATACATTGTGTCAAATAATG	2760			
OY		2761	AAACGAGAGAGTATGCTCATAGATTTTGGAATTACAGATGCAAAAGTAATAAGAAAG	2820			
Db		2761	AAACGAGAGAGTATGCTCATAGATTTTGGAATTACAGATGCAAAAGTAATAAGAAAG	2820			
OY		2821	AAAAAGAAAACCAGAGAAATATTTGCCAAGCACTTACAAGATTAACCCATCACTAATACT	2880			
Db		2821	AAAAAGAAAACCAGAGAAATATTTGCCAAGCACTTACAAGATTAACCCATCACTAATACT	2880			
OY		2881	TTCCTTTGCAAAAACTGCAGTGCTAGCCTGTCTGGGGAAGATATCCATGTAATTGAG	2940			
Db		2881	TTCCTTTGCAAAAACTGCAGTGCTAGCCTGTCTGGGGAAGATATCCATGTAATTGAG	2940			
OY		2941	AAATGATCATCAGTCAATATGACCCCAAGAAATCAAGAACTTTACATTGTAAGAAAAAC	3000			
Db		2941	AAATGATCATCAGTCAATATGACCCCAAGAAATCAAGAACTTTACATTGTAAGAAAAAC	3000			
OY		3001	AAAGCACTGC AAAAGAGTGTGCCGACTATCAATAAATGTTGTAATCATCTGCAAAATGT	3060			
Db		3001	AAAGCACTGC AAAAGAGTGTGCCGACTATCAATAAATGTTGTAATCATCTGCAAAATGT	3060			
OY		3061	GGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGGCTGTCTCAAAATA	3120			
Db		3061	GGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGGCTGTCTCAAAATA	3120			
OY		3121	AGGAATTTTGTAGTGTCTTCAAAAAATAATCAACAAAGAAACAATACAAAAAGTGGGTA	3180			
Db		3121	AGGAATTTTGTAGTGTCTTCAAAAAATAATCAACAAAGAAACAATACAAAAAGTGGGTA	3180			
OY		3181	GAATTACCTATCACATTTTCCCACATCTTGACTATTCCAAGATGCTGTTATTAGTGATGAG	3240			
Db		3181	GAATTACCTATCACATTTTCCCACATCTTGACTATTCCAAGATGCTGTTATTAGTGATGAG	3240			
OY		3241	GATTAGCACTTGATGAAGATCTTTTAAAAATACTATCAGTTAAACATTTAATATGATTA	3300			
Db		3241	GATTAGCACTTGATGAAGATCTTTTAAAAATACTATCAGTTAAACATTTAATATGATTA	3300			
OY		3301	TGATTAATGTATTCATTATGCTACAGAACTGCACATAAGAATCAATAAAATGATGTTTTTA	3360			
Db		3301	TGATTAATGTATTCATTATGCTACAGAACTGCACATAAGAATCAATAAAATGATGTTTTTA	3360			
OY		3361	CTCTG 3365				
Db		3361	CTCTG 3365				
<b>RESULT 3</b>							
ID	ADR14404	standard; DNA; 3380 BP.					
XX	ADR14404;						
DT	21-OCT-2004	(first entry)					
DE	Human NF-kappaB pathway-associated gene SeqID405.						
KW	NF-kappaB pathway; antiinflammatory; cyclostatic; hepatotropic; virucide;						
KW	antiarteriosclerotic; antineuritic; immunomodulator; gastrointestinal; vasotrophic;						
KW	immunorepressive; vulnerable; gene therapy; immune disorder;						
KW	inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;						
KW	hepatitis B; hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;						
KW	hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;						
KW	X-linked anhidrotic ectodermal dysplasia; immunodeficiency;						
KW	viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;						

KM viral replication; host cell survival; evasion of immune response;  
KM rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;  
KM atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;  
KM autoimmune disorder; hyper immune activity;  
KM aberrant acute phase response; hypercongenital condition; birth defect;  
KM necrotic lesion; wound; organ transplant rejection;  
KM aberrant signal transduction; proliferating disorder; cancer;  
KM HIV propagation; gene; ds; human.  
OS Homo sapiens.  
XX WO2004065577-A2.  
PN  
XX  
XX 05-AUG-2004.  
PD  
XX 13-JAN-2004; 2004WO-US000798.  
PF  
XX  
XX 14-JAN-2003; 2003US-0440068P.  
PR 12-MAY-2003; 2003US-0469757P.  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
PA  
XX  
XX Nadler SG, Neubauer MG, Feder JN, Carman J;  
PI  
XX WPI; 2004-562168/54.  
DR P-PSDB; ADR14405.  
XX  
XX  
XX New isolated polynucleotides and polypeptides associated with NF-kappab  
PT pathway, useful for diagnosing, treating, or preventing disorders or  
PT diseases associated with NF-kappab pathway.  
XX  
XX Claim 1; SEQ ID NO 405; 237pp; English.  
XX  
CC This invention relates to the novel association of protein sequences (and  
CC the genes which encode them) to the NF-kappab pathway. The invention may  
CC be useful for the production of compounds with an antiinflammatory,  
CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,  
CC gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,  
CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or  
CC vulnerary activity or for gene therapy. The proteins and nucleotides are  
CC useful for diagnosing, preventing, treating, or ameliorating conditions  
CC or diseases associated with the NF-kappab pathway. The condition is an  
CC immune disorder, an inflammatory disorder, an inflammatory disorder  
CC related to aberrant NF-kappab regulation, cancer, aberrant apoptosis,  
CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-Igm  
CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic  
CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,  
CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell  
CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory  
CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick  
CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper  
CC immune activity, disorders related to aberrant acute phase responses,  
CC hypercongenital conditions, birth defects, necrotic lesions, wounds,  
CC organ transplant rejection, conditions related to organ transplant  
CC rejection, disorders related to aberrant signal transduction,  
CC proliferating disorders, cancers and HIV propagation in cells infected  
CC with other viruses. The present sequence is that of a human gene which is  
CC subject to the novel association with the NF-kappab pathway of the  
CC invention. Note: This sequence does not appear in the specification but  
CC was obtained by the indexer from Genbank.  
XX  
SQ Sequence 3380 BP; 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Query Match 100.0%; Score 3365; DB 13; Length 3380;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCGCGGCTGAGAGCCCTGTGACAACCTGTCATTGTTCAGGACAGAGCGGTAGAC 60  
Db 1 GCGCGCGGCTGAGAGCCCTGTGACAACCTGTCATTGTTCAGGACAGAGCGGTAGAC 60  
QY 61 CTGCTTCTTAAGTGGCAGCGGACAGCGGACGACATTTCACCTGTCCCGACAGACAA 120

Db 61 CTGCTTCTTAAGTGGCAGCGGACAGCGGACGACGACATTTCACTGTCCCGACAGACAA 120  
QY 121 CAGCACCATCTGCTTGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTGAAATGGC 180  
Db 121 CAGCACCATCTGCTTGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTGAAATGGC 180  
QY 181 TATTCACAGACGAGAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAATAATG 240  
Db 181 TATTCACAGACGAGAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAATAATG 240  
QY 241 TACATCCAGGTGAGCCCTGTGCTGACTAATCTGACCTTCTGTGCTTCAGAGGTGAAGAG 300  
Db 241 TACATCCAGGTGAGCCCTGTGCTGACTAATCTGACCTTCTGTGCTTCAGAGGTGAAGAG 300  
QY 301 CAGATTCAGAGGACAGTCCGCCACCTCCGGGAACATGACAGGCAGTTGAACTGTGCTGAGC 360  
Db 301 CAGATTCAGAGGACAGTCCGCCACCTCCGGGAACATGACAGGCAGTTGAACTGTGCTGAGC 360  
QY 361 ACCTTGGAGAGGAGTCTTGACACCTTGTTGGAAGTTCGGGAAATTCGTGAGGCCCTCCGG 420  
Db 361 ACCTTGGAGAGGAGTCTTGACACCTTGTTGGAAGTTCGGGAAATTCGTGAGGCCCTCCGG 420  
QY 421 AGAACCGGACGCCCTCTGGCCGCCGCTACATGAAACCCCTGAGCTCACGGACTTGCCCTCT 480  
Db 421 AGAACCGGACGCCCTCTGGCCGCCGCTACATGAAACCCCTGAGCTCACGGACTTGCCCTCT 480  
QY 481 CCATCGTTTGAGAACGCTCATGATGATAATCTCCAACTGCTGAACCTCCTTCAGCCCACT 540  
Db 481 CCATCGTTTGAGAACGCTCATGATGATAATCTCCAACTGCTGAACCTCCTTCAGCCCACT 540  
QY 541 CTGTGGACAAGCTTCTAGTTAGAGAGTCTTGATAAGTGAATGAGAGGAACCTGTTG 600  
Db 541 CTGTGGACAAGCTTCTAGTTAGAGAGTCTTGATAAGTGAATGAGAGGAACCTGTTG 600  
QY 601 ACAATTGAAGACAGAAACCGGATGCTGCTGCAGAAAACAATGGAATGAATCAGGTGTA 660  
Db 601 ACAATTGAAGACAGAAACCGGATGCTGCTGCAGAAAACAATGGAATGAATCAGGTGTA 660  
QY 661 AGAGAGCTACTAAAGAGATTTGTGAGAAAGAAACCTGTTCTCTGATTTCTGAATGTT 720  
Db 661 AGAGAGCTACTAAAGAGATTTGTGAGAGTCTTGATTAAGTGAATGAGAGGAACCTGTTG 720  
QY 721 CTTCGTCAAACAGGAAACAATGAATGTAATGTTCCAAAGATTAAACAGGCTTGATGCTCAGAA 780  
Db 721 CTTCGTCAAACAGGAAACAATGAATGTAATGTTCCAAAGATTAAACAGGCTTGATGCTCAGAA 780  
QY 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGTTCTCAAGTGAAGAGCAACTT 840  
Db 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGTTCTCAAGTGAAGAGCAACTT 840  
QY 841 CTTTCAACCAAGTTTGCAGATTCTTCTAGTTTCAAGAAATGAGACACAAGTTTGGCAGAA 900  
Db 841 CTTTCAACCAAGTTTGCAGAAATCTGAGAAAGAGTCTGGGCGCATGAGAAATTAATCA 900  
QY 901 TCAGAAATCATTTTGCAGATTCTTCTAGTTTCAAGAAATGAGACACAAGTTTGGCAGAA 960  
Db 901 TCAGAAATCATTTTGCAGATTCTTCTAGTTTCAAGAAATGAGACACAAGTTTGGCAGAA 960  
QY 961 GGAAGTGTCAAGTCTTAGATGAAGTCTTGAGACATAACAGCAACATGGGCAATGATTC 1020  
Db 961 GGAAGTGTCAAGTCTTAGATGAAGTCTTGAGACATAACAGCAACATGGGCAATGATTC 1020  
QY 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA 1080  
Db 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA 1080  
QY 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTGGAAGGGAAGAAAT 1140  
Db 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTGGAAGGGAAGAAAT 1140  
QY 1141 ATCATCATCTGCTCCCTACAGGAGTGAAGAAACAGAGTGGCTGTTTACATTTGCCAAG 1200  
Db 1141 ATCATCATCTGCTCCCTACAGGAGTGAAGAAACAGAGTGGCTGTTTACATTTGCCAAG 1200

QY	1201	GATCACTTAGACAAGAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAT	1260
Db	1201	GATCACTTAGACAAGAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAT	1260
QY	1261	AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACTTTTGAAGAAATGG	1320
Db	1261	AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACTTTTGAAGAAATGG	1320
QY	1321	TATCGTGTATTGGATTAAAGTGTGATACCACACTGAAAAATATCATTTTCCAGAAGTTGTC	1380
Db	1321	TATCGTGTATTGGATTAAAGTGTGATACCACACTGAAAAATATCATTTTCCAGAAGTTGTC	1380
QY	1381	AAGTCCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAAACTTG	1440
Db	1381	AAGTCCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAAACTTG	1440
QY	1441	GAAAAATGAGAGAAGATGCTGTGTTCAATTGTGCAGACTTTTCCCTCATTTATCATTTGATGAA	1500
Db	1441	GAAAAATGAGAGAAGATGCTGTGTTCAATTGTGCAGACTTTTCCCTCATTTATCATTTGATGAA	1500
QY	1501	TGTCATCACACCAACAAGAAGACAGTGTAATAACATCATGAGGCATTATTGTATGCAG	1560
Db	1501	TGTCATCACACCAACAAGAAGACAGTGTAATAACATCATGAGGCATTATTGTATGCAG	1560
QY	1561	AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATCCCTTCTCAGATA	1620
Db	1561	AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATCCCTTCTCAGATA	1620
QY	1621	CTGGACTAACAGCTTCACTGCTGTGTGGAGGGCCACGAAGCAAGCCAAAGCTGAAGAA	1680
Db	1621	CTGGACTAACAGCTTCACTGCTGTGTGGAGGGCCACGAAGCAAGCCAAAGCTGAAGAA	1680
QY	1681	CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTTAATAAACTGTTAAGAAGAAAC	1740
Db	1681	CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTTAATAAACTGTTAAGAAGAAAC	1740
QY	1741	CTTGATCAACTGAAAAACCAATAACAGAGCCATGCAAGAACTTGGCCATTGCAGATGCA	1800
Db	1741	CTTGATCAACTGAAAAACCAATAACAGAGCCATGCAAGAACTTGGCCATTGCAGATGCA	1800
QY	1801	ACCAGAGAAGATCCATTTAAGAGAAACTCTAGAATAATGACAAGATTCAAACTTAAT	1860
Db	1801	ACCAGAGAAGATCCATTTAAGAGAAACTCTAGAATAATGACAAGATTCAAACTTAAT	1860
QY	1861	TGTCAAATGAGTCCAAATGTCAAGATTTTGGAACTCAACCTATGAACAATGGGCCATTCAA	1920
Db	1861	TGTCAAATGAGTCCAAATGTCAAGATTTTGGAACTCAACCTATGAACAATGGGCCATTCAA	1920
QY	1921	ATGGAATAAAAGCTGCATAAAAAAGGAATCGCAAGAAGCTGTTGTGCAGAACATTTG	1980
Db	1921	ATGGAATAAAAGCTGCATAAAAAAGGAATCGCAAGAAGCTGTTGTGCAGAACATTTG	1980
QY	1981	AGGAAGTACAAATGAGGCCCTTAACAATTAATGACACAATTCGAATGATAGTGCATTAAT	2040
Db	1981	AGGAAGTACAAATGAGGCCCTTAACAATTAATGACACAATTCGAATGATAGTGCATTAAT	2040
QY	2041	CATCTTGAACCTTCTATATAAGAAAGAAAGATTAAGAGTTTGCAGTCATAGAAAGATGAT	2100
Db	2041	CATCTTGAACCTTCTATATAAGAAAGAAAGATTAAGAGTTTGCAGTCATAGAAAGATGAT	2100
QY	2101	AGTGATGAGGAGTGATGATGATGATTTGTGATGGTGAAGATGAGATGATTTTAAAG	2160
Db	2101	AGTGATGAGGAGTGATGATGATGATTTGTGATGGTGAAGATGAGATGATTTTAAAG	2160
QY	2161	AAACCTTTGAACCTGATGAAAACAGATAGATTTCTCATGACTTTATTTTGAACAAT	2220
Db	2161	AAACCTTTGAACCTGATGAAAACAGATAGATTTCTCATGACTTTATTTTGAACAAT	2220
QY	2221	AAAAATGTTGAAAAAGGCTGGCTGAAAAACCAGAATATGAAAAATGAAAAAGCTGACCAATTA	2280
Db	2221	AAAAATGTTGAAAAAGGCTGGCTGAAAAACCAGAATATGAAAAATGAAAAAGCTGACCAATTA	2280

QY	2281	AGAAATACCATAATGAGCAATATATCTAGAGCTGAGGAATCAGCACGAGGAATTAATCTTT	2340
Db	2281	AGAAATACCATAATGAGCAATATATCTAGAGCTGAGGAATCAGCACGAGGAATTAATCTTT	2340
QY	2341	ACAAAAACACGACAGAGTGCATATGCGCTTCCAGTGGATTACTGAAAAATGAAAAATTT	2400
Db	2341	ACAAAAACACGACAGAGTGCATATGCGCTTCCAGTGGATTACTGAAAAATGAAAAATTT	2400
QY	2401	GCTGAAGTAGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAATTCAAA	2460
Db	2401	GCTGAAGTAGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAATTCAAA	2460
QY	2461	CCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAATTTCCGCACTGGAAAAATCAAT	2520
Db	2461	CCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAATTTCCGCACTGGAAAAATCAAT	2520
QY	2521	CTGCTTATCGCTACACAGTGGCAGAGAAGGCTGGATTATTTAAAGAAATGTAACTTTGTT	2580
Db	2521	CTGCTTATCGCTACACAGTGGCAGAGAAGGCTGGATTATTTAAAGAAATGTAACTTTGTT	2580
QY	2581	ATCCGTTATGCTCTCGTCAACCAATGAATAAGCATGGTCCAGGCCGTGTGAGCCAGA	2640
Db	2581	ATCCGTTATGCTCTCGTCAACCAATGAATAAGCATGGTCCAGGCCGTGTGAGCCAGA	2640
QY	2641	GCTGATGAGAGCACCTTAAGTCTCTGTTGCTCACAAGTGTTCAGAGATTATCGAACATGAG	2700
Db	2641	GCTGATGAGAGCACCTTAAGTCTCTGTTGCTCACAAGTGTTCAGAGATTATCGAACATGAG	2700
QY	2701	ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAGCTATACATTGTGTCAAAATATG	2760
Db	2701	ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAGCTATACATTGTGTCAAAATATG	2760
QY	2761	AAACGAGAGAGTATGCTCATTAAGATTTTGAATTACAGATGCCAAAGTATAATGAAAAAG	2820
Db	2761	AAACGAGAGAGTATGCTCATTAAGATTTTGAATTACAGATGCCAAAGTATAATGAAAAAG	2820
QY	2821	AAAATGAAAAACCAAGAGAAATATTGCCAAGCATTAACAAGATTAACCATCACTAATTAAT	2880
Db	2821	AAAATGAAAAACCAAGAGAAATATTGCCAAGCATTAACAAGATTAACCATCACTAATTAAT	2880
QY	2881	TTCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTCTGGGGAAGATATCCATGTAAATTGAG	2940
Db	2881	TTCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTCTGGGGAAGATATCCATGTAAATTGAG	2940
QY	2941	AAAATGCATCAGCTCAATATGACCCCAAGATTCAAGGAATTTACATTGTAAAGAAAAAC	3000
Db	2941	AAAATGCATCAGCTCAATATGACCCCAAGATTCAAGGAATTTACATTGTAAAGAAAAAC	3000
QY	3001	AAAGCACTGCAAAAAGAAAGTGTGCCGACTATCAATAAATGTTGAATCATCTGCAAAATGT	3060
Db	3001	AAAGCACTGCAAAAAGAAAGTGTGCCGACTATCAATAAATGTTGAATCATCTGCAAAATGT	3060
QY	3061	GGCCAGGCTTGGGGAACAATGATGTGCACAAAGCCTTAGATTTGCCCTGTCTCAAAAAATA	3120
Db	3061	GGCCAGGCTTGGGGAACAATGATGTGCACAAAGCCTTAGATTTGCCCTGTCTCAAAAAATA	3120
QY	3121	AGGAATTTTGTAGTGGTTTTCAAAAATAATTCAACAAGAACAATACAAAAAGTGGGTA	3180
Db	3121	AGGAATTTTGTAGTGGTTTTCAAAAATAATTCAACAAGAACAATACAAAAAGTGGGTA	3180
QY	3181	GAATTACCTATCACAATTTCCCAATCTTGACTAATTCAGAATGCTGTTATTTAGTGAATGAG	3240
Db	3181	GAATTACCTATCACAATTTCCCAATCTTGACTAATTCAGAATGCTGTTATTTAGTGAATGAG	3240
QY	3241	GATTAGACCTTGATTGAAGATTTCTTTAAAAATACTATCAGTTAAACATTTAAATATGATTA	3300
Db	3241	GATTAGACCTTGATTGAAGATTTCTTTAAAAATACTATCAGTTAAACATTTAAATATGATTA	3300
QY	3301	TGATTAATGTATTCATTAATGCTACAGAACTGCATTAAGAATCAATAAAATGATTGTTTAA	3360
Db	3301	TGATTAATGTATTCATTAATGCTACAGAACTGCATTAAGAATCAATAAAATGATTGTTTAA	3360
QY	3361	CTCTG 3365	

Db 3361 CTCTG 3365

RESULT 4  
AEA23691  
ID AEA23691 standard; DNA; 3380 BP.

XX AEA23691;  
AC  
XX  
DT 11-AUG-2005 (first entry)

XX Human PRO polypeptide DNA SEQ ID NO 233.

XX Immune disorder; PRO; Antinflammatory; Dermatological;  
KW Immunosuppressive; Antiarthritic; Osteopathic;  
KW Muscular-Gen.; Vasotropic; Antianemic; Antipsoriatic; Immunostimulant;  
KW ds; gene.

XX Homo sapiens.

XX WO2005051988-A2.

XX 09-JUN-2005.

XX 02-MAR-2004; 2004WO-US006460.

XX 03-MAR-2003; 2003US-0451884P.

XX (GETH ) GENENTECH INC.

PI Abbas A, Bodary S, Clark H, Schoenfeld J, Williams PM, Wood WI;  
PI Wu TD;

XX WPI; 2005-417958/42.  
DR P-PSDB; AEA23692.

XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and  
PT treating an immune related disorder, e.g. systemic lupus erythematosus,  
PT rheumatoid arthritis, osteoarthritis, autoimmune hemolytic anemia, or  
PT psoriasis.

XX Disclosure; SEQ ID NO 233; 966bp; English.

XX The invention relates to an isolated nucleic acid. The polypeptide,  
CC compound or composition, and methods are useful for diagnosing and  
CC treating an immune related disorder, e.g. systemic lupus erythematosus,  
CC rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis,  
CC spondyloarthropathies, systemic sclerosis, idiopathic inflammatory  
CC myopathies, Sjogren's syndrome, systemic vasculitis, sarcoidosis,  
CC autoimmune hemolytic anemia, autoimmune or immune-mediated skin diseases  
CC including bullous skin diseases, erythema multiforme and contact  
CC dermatitis, psoriasis, lymphadenopathy, splenomegaly and leukopenia. The  
CC present sequence represents a human PRO polypeptide DNA.

XX Sequence 3380 BP; 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Query Match 100.0%; Score 3365; DB 14; Length 3380;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCGCGGCGCTGAGAGCCCTGTGACAACTCGTCATTGTCAAGCACAGAGCGGTAGAC 60  
Db 1 GCGCGCGGCGCTGAGAGCCCTGTGACAACTCGTCATTGTCAAGCACAGAGCGGTAGAC 60  
QY 61 CCTGCTTCTTAAGTGGGCGAGCGGACGCGGACGACATTTCACTGTCCCGCAGACAA 120  
Db 61 CCTGCTTCTTAAGTGGGCGAGCGGACGCGGACGACATTTCACTGTCCCGCAGACAA 120  
QY 121 CAGCACCATCTGCTTGGGAGAACCCCTCTCTCTCTGAGAAAGAAAGATGTGAATGGG 180  
Db 121 CAGCACCATCTGCTTGGGAGAACCCCTCTCTCTCTGAGAAAGAAAGATGTGAATGGG 180

QY 181 TATTCACAGACGAGAATTTCCGCTATCTCATCTCGTCTTCAAGGCCAGGGTGAAATG 240  
Db 181 TATTCACAGACGAGAATTTCCGCTATCTCATCTCGTCTTCAAGGCCAGGGTGAAATG 240  
QY 241 TACATCCAGGTGAGCCCTGTGCTGACTACCTGACCTTTCTGCTTGACAGAGGTGAAGAG 300  
Db 241 TACATCCAGGTGAGCCCTGTGCTGACTACCTGACCTTTCTGCTTGACAGAGGTGAAGAG 300  
QY 301 CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAAGGCACTGAACTGCTGTGACC 360  
Db 301 CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAAGGCACTGAACTGCTGTGACC 360  
QY 361 ACCTTGAGAGGAGTCTGGCACCTTGTTGACTCGGGAATTCGTGAGGCCCTCCGG 420  
Db 361 ACCTTGAGAGGAGTCTGGCACCTTGTTGACTCGGGAATTCGTGAGGCCCTCCGG 420  
QY 421 AGAACCGGAGCCCTCTGGCCCGGCTACATGAACCTGAGCTCAGGACTTGCCCTCT 480  
Db 421 AGAACCGGAGCCCTCTGGCCCGGCTACATGAACCTGAGCTCAGGACTTGCCCTCT 480  
QY 481 CCATCGTTGAGACGCTCATGATGATATCTCCACTGCTGAACCTCTCAGCCCACT 540  
Db 481 CCATCGTTGAGACGCTCATGATGATATCTCCACTGCTGAACCTCTCAGCCCACT 540  
QY 541 CTGTGACACAGCTTCTAGTTAGACGCTCTTGATAGTGATGAGAGGAACTGTG 600  
Db 541 CTGTGACACAGCTTCTAGTTAGACGCTCTTGATAGTGATGAGAGGAACTGTG 600  
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAACAAATGAATGAATCAGGTGA 660  
Db 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAACAAATGAATGAATCAGGTGA 660  
QY 661 AGAGAGCTACTAAAAAGGATTGTGAGAAAGAAACTGTTCTCTGATTCTGAATGTT 720  
Db 661 AGAGAGCTACTAAAAAGGATTGTGAGAAAGAAACTGTTCTCTGATTCTGAATGTT 720  
QY 721 CTTGCTCAACAGGAAACAAATGAACCTGTCCAAGAGTTAAAGGCTCTGATTGCTCAGAA 780  
Db 721 CTTGCTCAACAGGAAACAAATGAACCTGTCCAAGAGTTAAAGGCTCTGATTGCTCAGAA 780  
QY 781 AGCAATGCAGAGATTGAGATTTATCACAAGTTGATGCTCTCAAGTGAAGCAACTT 840  
Db 781 AGCAATGCAGAGATTGAGATTTATCACAAGTTGATGCTCTCAAGTGAAGCAACTT 840  
QY 841 CTTCAACCAAGTTCAAGCCAAATCTGAGAGAGAGGTCTGGGCGCATGGAATACTCA 900  
Db 841 CTTCAACCAAGTTCAAGCCAAATCTGAGAGAGAGGTCTGGGCGCATGGAATACTCA 900  
QY 901 TCAGAAATCATCTTTGACAGATTCCTGTAGTTTCAGAAATCAGACACAAGTTGGCAGAA 960  
Db 901 TCAGAAATCATCTTTGACAGATTCCTGTAGTTTCAGAAATCAGACACAAGTTGGCAGAA 960  
QY 961 GGAAGTGCAGCTGCTTAGATGAAGTCTTGACATTAACAGCAACATGGGCACTGATCA 1020  
Db 961 GGAAGTGCAGCTGCTTAGATGAAGTCTTGACATTAACAGCAACATGGGCACTGATCA 1020  
QY 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAAATGTGCGACGAAAGACATCCCCGAGCCA 1080  
Db 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAAATGTGCGACGAAAGACATCCCCGAGCCA 1080  
QY 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCTTGGAAGGGAAGAT 1140  
Db 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCTTGGAAGGGAAGAT 1140  
QY 1141 ATCATCATCTGCTCTCTACAGGAGTGAAGAAACAGAGTGTGTTACATTGCCAAG 1200  
Db 1141 ATCATCATCTGCTCTCTACAGGAGTGAAGAAACAGAGTGTGTTACATTGCCAAG 1200  
QY 1201 GATCACTTAGACAAAGAAAAAGCACTGAGCCTGGAAGTTATAGTCTTGTCAAT 1260  
Db 1201 GATCACTTAGACAAAGAAAAAGCACTGAGCCTGGAAGTTATAGTCTTGTCAAT 1260  
QY 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCAACATTTTGAAGAAATGG 1320

Db 1261 AAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAACCATTTTGAAGAAATGG 1320  
QY 1321 TATCGTGTATTGATTAAGTGTGATACCCCACTGAAATATCATTTCCAGAAGTTGTC 1380  
Db 1321 TATCGTGTATTGATTAAGTGTGATACCCCACTGAAATATCATTTCCAGAAGTTGTC 1380  
QY 1381 AAGTCTGTGATATTATTCAGTACAGCTCAATCCTTGAAAACCTCCCTTTAACTTG 1440  
Db 1381 AAGTCTGTGATATTATTCAGTACAGCTCAATCCTTGAAAACCTCCCTTTAACTTG 1440  
QY 1441 GAAAATGAGAGAGATGCTGCTGTTCAATTGTTCAGACTTTTCCCTCATTTATCATTTGAA 1500  
Db 1441 GAAAATGAGAGAGATGCTGCTGTTCAATTGTTCAGACTTTTCCCTCATTTATCATTTGAA 1500  
QY 1501 TGTTCATCACACCAACAAAGAGCAGTGTATTAATACATCATGAGGCATTAATTGATGAG 1560  
Db 1501 TGTTCATCACACCAACAAAGAGCAGTGTATTAATACATCATGAGGCATTAATTGATGAG 1560  
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACAGTGATTCCTCCCTCAGATA 1620  
Db 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACAGTGATTCCTCCCTCAGATA 1620  
QY 1621 CTGGGACTAACAGCTTCACTGCTGTTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
Db 1621 CTGGGACTAACAGCTTCACTGCTGTTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
QY 1681 CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACTGTGTTAAAGAAAC 1740  
Db 1681 CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACTGTGTTAAAGAAAC 1740  
QY 1741 CTGTATCAACTGAAAAACAATACAGAGCCATGCAAGAGTTGCCATTCAGATGCA 1800  
Db 1741 CTGTATCAACTGAAAAACAATACAGAGCCATGCAAGAGTTGCCATTCAGATGCA 1800  
QY 1801 ACCAGAGAAGATCCATTTAAAGAGAACTTCTGAATAATGACAAGAGTTCAAACTTAT 1860  
Db 1801 ACCAGAGAAGATCCATTTAAAGAGAACTTCTGAATAATGACAAGAGTTCAAACTTAT 1860  
QY 1861 TGTCAATAGAGTCCCAATGTGAGATTTTGGAACTCAACCTATGAAACAATGGGCCATTCAA 1920  
Db 1861 TGTCAATAGAGTCCCAATGTGAGATTTTGGAACTCAACCTATGAAACAATGGGCCATTCAA 1920  
QY 1921 ATGGAAGAAAAAGCTGCAAAAAAGAAATCGCAAGAAAGCTGTTGTGCAAGAACTTTG 1980  
Db 1921 ATGGAAGAAAAAGCTGCAAAAAAGAAATCGCAAGAAAGCTGTTGTGCAAGAACTTTG 1980  
QY 1981 AGGAAGTACAATGAGGCCCTCAAAATTAATGACAACAATTCGAATGATAGTGCCTACT 2040  
Db 1981 AGGAAGTACAATGAGGCCCTCAAAATTAATGACAACAATTCGAATGATAGTGCCTACT 2040  
QY 2041 CATCTTGAACCTTCTAATAAGAGAGAAAGATGAAGTTGCAGTCAATAGAGATGAT 2100  
Db 2041 CATCTTGAACCTTCTAATAAGAGAGAAAGATGAAGTTGCAGTCAATAGAGATGAT 2100  
QY 2101 AGTGATGAGGTTGTGATGATGAGTATTTGTGATGCTGATGAAGATGAGATTTAAAG 2160  
Db 2101 AGTGATGAGGTTGTGATGATGAGTATTTGTGATGCTGATGAAGATGAGATTTAAAG 2160  
QY 2161 AAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGAACAAT 2220  
Db 2161 AAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGAACAAT 2220  
QY 2221 AAAATGTTGAAAAAGCTGCTGAAAAACCCAGAAATGAAAAAGCTGACCAAAATTA 2280  
Db 2221 AAAATGTTGAAAAAGCTGCTGAAAAACCCAGAAATGAAAAAGCTGACCAAAATTA 2280  
QY 2281 AGAATAACCATATAGAGCAATATACTAGACTGAGGAATCAGCAGAGAAATAATCTTT 2340  
Db 2281 AGAATAACCATATAGAGCAATATACTAGACTGAGGAATCAGCAGAGAAATAATCTTT 2340  
QY 2341 ACAAAAAACAGCAGAGTGCATATGCGCTTCCCACTGATTACTGAAAAATGAAAAATTT 2400  
Db 2341 ACAAAAAACAGCAGAGTGCATATGCGCTTCCCACTGATTACTGAAAAATGAAAAATTT 2400

Db 2341 ACAAAAAACAGCAGAGTGCATATGCGCTTCCCACTGATTACTGAAAAATGAAAAATTT 2400  
QY 2401 GCTGAAGTAGAGTCAAGCCCAACCATCTGATTTGAGCTGACACAGCAGTGAATTCAA 2460  
Db 2401 GCTGAAGTAGAGTCAAGCCCAACCATCTGATTTGAGCTGACACAGCAGTGAATTCAA 2460  
QY 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAATTTCCGACTGGAAAAATCAAT 2520  
Db 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAATTTCCGACTGGAAAAATCAAT 2520  
QY 2521 CTGCTTATCGCTACCAAGTGGCAGAGAGAGTCTGATATTAAAGATGTAACATTGTT 2580  
Db 2521 CTGCTTATCGCTACCAAGTGGCAGAGAGAGTCTGATATTAAAGATGTAACATTGTT 2580  
QY 2581 ATCCGTTATGCTCTGTCACCAATGAATAGCCATGTCAGAGCCCGCTGTCAGCCAGA 2640  
Db 2581 ATCCGTTATGCTCTGTCACCAATGAATAGCCATGTCAGAGCCCGCTGTCAGCCAGA 2640  
QY 2641 GCTGATGAGAGCACTTACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700  
Db 2641 GCTGATGAGAGCACTTACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700  
QY 2701 ACAGTTAATGATTTCCGAGAGAGATGATGTAATAAGCTATACATTTGTTCAAAATATG 2760  
Db 2701 ACAGTTAATGATTTCCGAGAGAGATGATGTAATAAGCTATACATTTGTTCAAAATATG 2760  
QY 2761 AAACAGAGAGATGCTCATTAAGATTTTGAATTAACAGATGCAAAAGTAAATGAAAAAG 2820  
Db 2761 AAACAGAGAGATGCTCATTAAGATTTTGAATTAACAGATGCAAAAGTAAATGAAAAAG 2820  
QY 2821 AAAATGAAAAACCAAGAGAAATATTTGCCAAGCTTACAAGAAATTAACCACTAATAACT 2880  
Db 2821 AAAATGAAAAACCAAGAGAAATATTTGCCAAGCTTACAAGAAATTAACCACTAATAACT 2880  
QY 2881 TTCTTTTGCAAAAACCTGCAAGTGTGCTAGCTGTTCTGGGAGAGATATTCATGTAATTGAG 2940  
Db 2881 TTCTTTTGCAAAAACCTGCAAGTGTGCTAGCTGTTCTGGGAGAGATATTCATGTAATTGAG 2940  
QY 2941 AAAATGCATCAGCTCAATATGACCCCAAGAAATTCAGAGAACTTTACATTTGAAGAAAAAC 3000  
Db 2941 AAAATGCATCAGCTCAATATGACCCCAAGAAATTCAGAGAACTTTACATTTGAAGAAAAAC 3000  
QY 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAATAAATGTTGATGCAAAATGT 3060  
Db 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAATAAATGTTGATGCAAAATGT 3060  
QY 3061 GGCCAGGCTTGGGGAACAATGATGTGCAAAAGGCTTAGATTTGCTGCTCAAAATA 3120  
Db 3061 GGCCAGGCTTGGGGAACAATGATGTGCAAAAGGCTTAGATTTGCTGCTCAAAATA 3120  
QY 3121 AGGAATTTTGTAGTGTGTTTCAAAAATATTCACAAAGAAACAATACAAAAAGTGCGTA 3180  
Db 3121 AGGAATTTTGTAGTGTGTTTCAAAAATATTCACAAAGAAACAATACAAAAAGTGCGTA 3180  
QY 3181 GAATTAACCTATCACTTCCCAATCTTGACTATTCAGAAATGCTGTTAATTAATGATGAG 3240  
Db 3181 GAATTAACCTATCACTTCCCAATCTTGACTATTCAGAAATGCTGTTAATTAATGATGAG 3240  
QY 3241 GATTAACCTTGAATGAGATTTCTTTAAATACTATCAGTTAAACATTAATATGATTA 3300  
Db 3241 GATTAACCTTGAATGAGATTTCTTTAAATACTATCAGTTAAACATTAATATGATTA 3300  
QY 3301 TGAATTAATGTATTCATTAATGCTACAGAACTGACATAAAGATCAATAAATGATTTGTTTA 3360  
Db 3301 TGAATTAATGTATTCATTAATGCTACAGAACTGACATAAAGATCAATAAATGATTTGTTTA 3360  
QY 3361 CTCTG 3365  
Db 3361 CTCTG 3365

RESULT 5  
AEA36113

ID AEA36113 standard; DNA; 3380 BP.  
XX  
AC AEA36113;  
XX  
DT 25-AUG-2005 (first entry)  
XX  
DE Human nucleic acid sequence #45.  
XX  
KW Screening; gene expression; colorectal tumor; colitis; Crohns disease;  
KW irritable bowel syndrome; gastrointestinal disease; cytostatic;  
KW gastrointestinal-gen.; antiinflammatory; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO2005054507-A2.  
XX  
PD 16-JUN-2005.  
XX  
PF 03-DEC-2004; 2004WO-GB005078.  
XX  
PR 04-DEC-2003; 2003GB-00028048.  
XX  
PA (UYSH-) UNIV SHEFFIELD.  
XX  
PI Corfe B, Chirakkal H;  
XX  
DR WPI; 2005-435407/44.  
XX  
XX  
PT Screening for nucleic acid molecules exhibiting altered expression in  
PT cells grown in the presence of butyrate, and detection of the nucleic  
PT acid molecules or the encoded polypeptides in diagnosing colorectal  
PT cancer.  
XX  
XX  
PS Disclosure; Page 77-78; 266pp; English.  
XX  
XX  
CC The invention relates to a method of screening for nucleic acid molecules  
CC that show altered expression in a first cell sample comprising comparing  
CC the gene expression profile of the sample with that of a second reference  
CC sample, where the first sample has been grown in the presence of butyrate  
CC or a related carbon source from which butyrate is directly or indirectly  
CC derived, but the reference sample has not. The invention also relates to  
CC a method of detecting at least one nucleic acid molecule associated with  
CC the initiation and/or progression of colorectal cancer in an animal,  
CC comprising providing a biological sample comprising at least one cell to  
CC be tested, contacting the sample with a ligand (preferably a hybridizing  
CC nucleic acid molecule) which binds to at least one nucleic acid and  
CC detecting the presence of at least one molecule in the sample, a method  
CC of detecting at least one polypeptide associated with the initiation  
CC and/or progression of colorectal cancer in an animal comprising providing  
CC a biological sample comprising at least one cell to be tested, contacting  
CC the sample with at least one ligand that specifically binds at least one  
CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
CC an amino acid sequence which varies by the addition, deletion or  
CC substitution of at least one amino acid residue and detecting the  
CC presence of the polypeptide in the sample, a method of screening for  
CC agents that modulate the activity of at least one polypeptide encoded by  
CC a gene associated with the initiation and/or progression of colorectal  
CC cancer comprising forming a preparation comprising at least one  
CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
CC at least one addition, deletion or substitution and at least one agent to  
CC be tested and determining the activity of the agent with respect to  
CC activity of the polypeptide, and an antibody or its effective binding  
CC portion identified by the method, for use as a pharmaceutical. The  
CC methods are useful for screening for nucleic acid molecules that show  
CC altered expression in a cell sample, and for detecting a nucleic acid and  
CC a polypeptide respectively, that are associated with the initiation  
CC and/or progression of colorectal cancer and are useful for detecting or  
CC monitoring colorectal cancer, especially adenocarcinoma. The methods are  
CC also useful for screening for agents that modulate the activity of at  
CC least one polypeptide encoded by a gene associated with the initiation  
CC and/or progression of cancer, where agents identified by the method are  
CC useful for treating colorectal cancer. The methods could also be used to  
CC detect or monitor other conditions such as colitis, Crohn's disease or

CC irritable bowel syndrome, as a screening tool for fiber consumption, as  
CC an assay for colon microflora functionality or for early detection of pre  
CC -cancerous growth. This sequence represents a human nucleic acid  
CC identified by the screening method of the invention.  
XX  
SQ Sequence 3380 BP; 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;  
Query Match 100.0%; Score 3365; DB 14; Length 3380;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGCGCCGCGCTGAGAGCCCTGTGACACCTCGTCATTGTGACGACAGCGGTAGAC 60  
DB 1 GCGCGCCGCGCTGAGAGCCCTGTGACACCTCGTCATTGTGACGACAGCGGTAGAC 60  
QY 61 CCTGCTTCTTAAGTGGGACGCGACGCGGACGACATTTCACTGTCCCGACAGCAA 120  
DB 61 CCTGCTTCTTAAGTGGGACGCGGACGCGGACGACATTTCACTGTCCCGACAGCAA 120  
QY 121 CAGCACCATCTGCTTGGAGAAACCTCTCCCTTCTGTAGAAAGAAAGATGTGAAATGG 180  
DB 121 CAGCACCATCTGCTTGGAGAAACCTCTCCCTTCTGTAGAAAGAAAGATGTGAAATGG 180  
QY 181 TATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTACGGCCAGGTAATATG 240  
DB 181 TATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTACGGCCAGGTAATATG 240  
QY 241 TACATCCAGGTGAGCCCTGTGCTGACCTACCTGACCTTCTGCTGAGAGGTGAAGAG 300  
DB 241 TACATCCAGGTGAGCCCTGTGCTGACCTACCTGACCTTCTGCTGAGAGGTGAAGAG 300  
QY 301 CAGATTGAGAGACAGTCCGCCACCTCCGGAAACATGACAGGCAATTGAATGCTGTGAGC 360  
DB 301 CAGATTGAGAGACAGTCCGCCACCTCCGGAAACATGACAGGCAATTGAATGCTGTGAGC 360  
QY 361 ACCTTGAGAAAGGAGTCTGGCACCTTGTTGAACTCGGGAATTCGTGAGGCCCTCCGG 420  
DB 361 ACCTTGAGAAAGGAGTCTGGCACCTTGTTGAACTCGGGAATTCGTGAGGCCCTCCGG 420  
QY 421 AGAACCGGAGCCCTCTGGCCGCCGCTACATGAACCCCTGAGCTACGGACTTGCCCTCT 480  
DB 421 AGAACCGGAGCCCTCTGGCCGCCGCTACATGAACCCCTGAGCTACGGACTTGCCCTCT 480  
QY 481 CCATCGTTGAGACGCTCATGATGAATATCTCCAAGTGTGACCTCTTACGCCCACT 540  
DB 481 CCATCGTTGAGACGCTCATGATGAATATCTCCAAGTGTGACCTCTTACGCCCACT 540  
QY 541 CTGGTGGACAGCTTCTAGTTAGAGCGTCTTGAGATAAGTGCATGAGAGAACTGTTG 600  
DB 541 CTGGTGGACAGCTTCTAGTTAGAGCGTCTTGAGATAAGTGCATGAGAGAACTGTTG 600  
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAACAATGAAATGAATCAGGTGTA 660  
DB 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAACAATGAAATGAATCAGGTGTA 660  
QY 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAACTGGTCTCTGCAATTTCTGAATGT 720  
DB 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAACTGGTCTCTGCAATTTCTGAATGT 720  
QY 721 CTTGCTCAACACAGAAACAATGAAGTTTCAAGAGTTAACAGGCTGTGATTGCTCAGAA 780  
DB 721 CTTGCTCAACACAGAAACAATGAAGTTTCAAGAGTTAACAGGCTGTGATTGCTCAGAA 780  
QY 781 AGCAATGCAGAGATTGAGAAATTTATCACAAAGTTGATGTCCTCAAGTGAAGACAATT 840  
DB 781 AGCAATGCAGAGATTGAGAAATTTATCACAAAGTTGATGTCCTCAAGTGAAGACAATT 840  
QY 841 CTTTCAACACAGATTGAGCCAAATCTGGAGAGAGGTCTGGGGCATGGAATTAAGTCA 900  
DB 841 CTTTCAACACAGATTGAGCCAAATCTGGAGAGAGGTCTGGGGCATGGAATTAAGTCA 900  
QY 901 TCAGAAATCATCTTTGAGATTCTTCTGATGTTTCAGAATCAGACCAAGTTGGCAGAA 960  
DB 901 TCAGAAATCATCTTTGAGATTCTTCTGATGTTTCAGAATCAGACCAAGTTGGCAGAA 960

Db 901 TCAGATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAA 960  
Qy 961 GGAAGTGTCAAGCTGCTTAGATGAAAAGTCTTGGACATTAACAGCAACATGGGCAGTATTCA 1020  
Db 961 GGAAGTGTCAAGCTGCTTAGATGAAAAGTCTTGGACATTAACAGCAACATGGGCAGTATTCA 1020  
Qy 1021 GGCAACATGGGAAGTGTATTCAGATGAAGAAATGGGCAGCAAGAGCATCCCCGAGCCA 1080  
Db 1021 GGCAACATGGGAAGTGTATTCAGATGAAGAAATGGGCAGCAAGAGCATCCCCGAGCCA 1080  
Qy 1081 GAACTCCAGCTCAGGCTTTACCAATGGAAGTTGCCAGCCAGCCTTGGAAAGGAAGAAAT 1140  
Db 1081 GAACTCCAGCTCAGGCTTTACCAATGGAAGTTGCCAGCCAGCCTTGGAAAGGAAGAAAT 1140  
Qy 1141 ATCATCATCTGCCCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTGGCCAAG 1200  
Db 1141 ATCATCATCTGCCCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTTTACATTGGCCAAG 1200  
Qy 1201 GATCACTTAGACAAGAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAT 1260  
Db 1201 GATCACTTAGACAAGAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAT 1260  
Qy 1261 AAGTAAGTGTAGTTGAACAGCTCTTCCGCAAGAGTTCACACCATTTTGAAGAAATGG 1320  
Db 1261 AAGTAAGTGTAGTTGAACAGCTCTTCCGCAAGAGTTCACACCATTTTGAAGAAATGG 1320  
Qy 1321 TATCGTGTATTGATTAAGTGGTGATACCCCACTGAAATATCATTTCCAGAAAGTTGTC 1380  
Db 1321 TATCGTGTATTGATTAAGTGGTGATACCCCACTGAAATATCATTTCCAGAAAGTTGTC 1380  
Qy 1381 AAGTCCCTGTGATATTATATCAGTACAGCTCAAAATCCTGAAAACTCCCTCTAAACTTG 1440  
Db 1381 AAGTCCCTGTGATATTATATCAGTACAGCTCAAAATCCTGAAAACTCCCTCTAAACTTG 1440  
Qy 1441 GAAAAATGGAAGAGATGCTGGTGTTCATTTGTCAAGCTTTTCCCTCATTTATCATTTGATGA 1500  
Db 1441 GAAAAATGGAAGAGATGCTGGTGTTCATTTGTCAAGCTTTTCCCTCATTTATCATTTGATGA 1500  
Qy 1501 TGTTCATCACACCAACAAAGAGCAGTGTATTAATTAATCATCATGAGGCATTAATTGATCAG 1560  
Db 1501 TGTTCATCACACCAACAAAGAGCAGTGTATTAATTAATCATCATGAGGCATTAATTGATCAG 1560  
Qy 1561 AAGTTGAAAAACAATAGACTCAAGAAAAAGAAAAACAACAGTGATTCCTCTCCAGATA 1620  
Db 1561 AAGTTGAAAAACAATAGACTCAAGAAAAAGAAAAACAACAGTGATTCCTCTCCAGATA 1620  
Qy 1621 CTGGGACTTAACAGCTTCACTGGTGTGGAGGGGCCAGAGCAAGCCAAAGCTGAAGAA 1680  
Db 1621 CTGGGACTTAACAGCTTCACTGGTGTGGAGGGGCCAGAGCAAGCCAAAGCTGAAGAA 1680  
Qy 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACTGTTAAAGAAAAAC 1740  
Db 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACTGTTAAAGAAAAAC 1740  
Qy 1741 CTTCATCAACTGAAAAACCAATAACAGAGCCATGCAAGAGTTGGCATTGACAGATGCA 1800  
Db 1741 CTTCATCAACTGAAAAACCAATAACAGAGCCATGCAAGAGTTGGCATTGACAGATGCA 1800  
Qy 1801 ACCAGAGAAGATCCATTAAAGAGAAACTTCTAGAAATATGACAAGAGATTCAAACTTAT 1860  
Db 1801 ACCAGAGAAGATCCATTAAAGAGAAACTTCTAGAAATATGACAAGAGATTCAAACTTAT 1860  
Qy 1861 TGTCAAAATGAGTCCAAATGTCAAGATTTTGGAACTCAACCTATGAACAATGGGCCATTCAA 1920  
Db 1861 TGTCAAAATGAGTCCAAATGTCAAGATTTTGGAACTCAACCTATGAACAATGGGCCATTCAA 1920  
Qy 1921 ATGGAAAAAAAAGCTGCAAAAAAGGAATCGCAAGAACGTGTTGTGCAAGAACATTGG 1980  
Db 1921 ATGGAAAAAAAAGCTGCAAAAAAGGAATCGCAAGAACGTGTTGTGCAAGAACATTGG 1980  
Qy 1981 AGGAAGTACAATGAGGCCCTTACAAATTAATGACACAAATTCGAATGATAGTGCCTATACT 2040  
Db 1981 AGGAAGTACAATGAGGCCCTTACAAATTAATGACACAAATTCGAATGATAGTGCCTATACT 2040

Qy 2041 CATCTGAAAACCTTTCTATTAATGAAGAGAAAGATTAAGAAAGTTTGCAGTCATAGAAGATGAT 2100  
Db 2041 CATCTGAAAACCTTTCTATTAATGAAGAGAAAGATTAAGAAAGTTTGCAGTCATAGAAGATGAT 2100  
Qy 2101 AGTGATGAGGGTGGTGATGATGAGTATTTGATGATGGTGATGAAGATGAGATGATTTAAAG 2160  
Db 2101 AGTGATGAGGGTGGTGATGATGAGTATTTGATGATGGTGATGAAGATGAGATGATTTAAAG 2160  
Qy 2161 AAACCTTTGAAACTGATGAAGAAACAGATAGATTCTCATGACTTTATTTTGAACAACAT 2220  
Db 2161 AAACCTTTGAAACTGATGAAGAAACAGATAGATTCTCATGACTTTATTTTGAACAACAT 2220  
Qy 2221 AAAATGTTGAAAAAGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAGCTGACCAATTA 2280  
Db 2221 AAAATGTTGAAAAAGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAGCTGACCAATTA 2280  
Qy 2281 AGAAATACCATTAATGAGCAATATATCTAGACTGAGGAATCAGCAGGAAATATCTTT 2340  
Db 2281 AGAAATACCATTAATGAGCAATATATCTAGACTGAGGAATCAGCAGGAAATATCTTT 2340  
Qy 2341 ACAAAAACACGACAGAGTGCATATATGCGCTTCCAGTGATTAATGAAAAATTTT 2400  
Db 2341 ACAAAAACACGACAGAGTGCATATATGCGCTTCCAGTGATTAATGAAAAATTTT 2400  
Qy 2401 GCTGAAGTGAAGTCAAAAGCCCACTCTGATTTGAGCTGGAACACAGCAGTGAAGTTCAA 2460  
Db 2401 GCTGAAGTGAAGTCAAAAGCCCACTCTGATTTGAGCTGGAACACAGCAGTGAAGTTCAA 2460  
Qy 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAATTTCCGACTGGAATAATCAAT 2520  
Db 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAATTTCCGACTGGAATAATCAAT 2520  
Qy 2521 CTGCTTATCGCTAACCAAGTGGCAGAAAGAGTCTGATATTTAAAGATGTAACATTGTT 2580  
Db 2521 CTGCTTATCGCTAACCAAGTGGCAGAAAGAGTCTGATATTTAAAGATGTAACATTGTT 2580  
Qy 2581 ATCCGTTATGTCCTGTCACCAATGAATAGCATGTCAGAGCCCGTGGTGAAGCCAGA 2640  
Db 2581 ATCCGTTATGTCCTGTCACCAATGAATAGCATGTCAGAGCCCGTGGTGAAGCCAGA 2640  
Qy 2641 GCTGATGAGAGCACCCTACGTCCTGTTGCTTCAAGTGGTTCAGAGATTATCGAACATGAG 2700  
Db 2641 GCTGATGAGAGCACCCTACGTCCTGTTGCTTCAAGTGGTTCAGAGATTATCGAACATGAG 2700  
Qy 2701 ACAGTTAATGATTTCCGAGAGAGATGATGATTAAGCTATACATTTGTTCAAAATATG 2760  
Db 2701 ACAGTTAATGATTTCCGAGAGAGATGATGATTAAGCTATACATTTGTTCAAAATATG 2760  
Qy 2761 AAACGAGAGAGTATGCTCATTAAGATTTTGAATTAACAGATGCAAGATTAATGAAAAAG 2820  
Db 2761 AAACGAGAGAGTATGCTCATTAAGATTTTGAATTAACAGATGCAAGATTAATGAAAAAG 2820  
Qy 2821 AAAATGAAGAACCAAGAGAAATATTTGCCAAGCATTAACAAGATTAACCATCAATAACT 2880  
Db 2821 AAAATGAAGAACCAAGAGAAATATTTGCCAAGCATTAACAAGATTAACCATCAATAACT 2880  
Qy 2881 TTCCTTTGCAAAAACTGCAAGTGTGCTAGCTGTTCTGGGGAAGATATTCATGTAATTGAG 2940  
Db 2881 TTCCTTTGCAAAAACTGCAAGTGTGCTAGCTGTTCTGGGGAAGATATTCATGTAATTGAG 2940  
Qy 2941 AAAATGCATCAGTCAATATGACCCCAAGATTTCAAGGAACCTTTACATTTGTAAGAGAAAAAC 3000  
Db 2941 AAAATGCATCAGTCAATATGACCCCAAGATTTCAAGGAACCTTTACATTTGTAAGAGAAAAAC 3000  
Qy 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATTAATGATGTAATTCATCTGCAAAATGT 3060  
Db 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATTAATGATGTAATTCATCTGCAAAATGT 3060  
Qy 3061 GGCCAGGCTTGGGGAACAATGATGTTGCAAAAGGCTTAGATTTGCTGTCTCAAAATA 3120  
Db 3061 GGCCAGGCTTGGGGAACAATGATGTTGCAAAAGGCTTAGATTTGCTGTCTCAAAATA 3120

QY 3121 AGAATTTTGTAGTGGTTTCAAAAATAATTCACAAAGAAACAATACAAAAGTGGTA 3180  
|||||  
Db 3121 AGGAATTTTGTAGTGGTTTCAAAAATAATTCACAAAGAAACAATACAAAAGTGGTA 3180  
QY 3181 GAATTAACCTATCACAATTTCCCAATCTTGACTATTCAGAATGCTGTTATTAGTATGAG 3240  
|||||  
Db 3181 GAATTAACCTATCACAATTTCCCAATCTTGACTATTCAGAATGCTGTTATTAGTATGAG 3240  
QY 3241 GATTAGCACTTGATTGAAGATTCTTTAAATACTATCAGTTAAACATTAAATATGATTA 3300  
|||||  
Db 3241 GATTAGCACTTGATTGAAGATTCTTTAAATACTATCAGTTAAACATTAAATATGATTA 3300  
QY 3301 TGATTAACTGATTATTCATTATGCTACAGAACTGACATTAAGAATCAATAAATGATTGTTTA 3360  
|||||  
Db 3301 TGATTAACTGATTATTCATTATGCTACAGAACTGACATTAAGAATCAATAAATGATTGTTTA 3360  
QY 3361 CTCTG 3365  
|||||  
Db 3361 CTCTG 3365

RESULT 6

ADY17563  
ID ADY17563 standard; DNA; 3434 BP.

XX  
AC ADY17563;

DT 05-MAY-2005 (first entry)

XX  
DE DNA encoding a PRO polypeptide, SEQ ID NO 3369.

XX  
KW Antinflammatory; Immune disorder; Dermatological; Immunosuppressive;  
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;  
KW Antihydroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;  
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;  
KW Antiallergic; ds; gene; diagnosis.

XX  
OS Homo sapiens.

XX  
PN WO2005016962-A2.

XX  
PD 24-FEB-2005.

XX  
PF 11-AUG-2004; 2004WO-US026249.

XX  
PR 11-AUG-2003; 2003US-0493546P.

XX  
PA (GETH ) GENENTECH INC.

PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;

DR WPI; 2005-182330/19.

XX  
PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and  
PT treating an immune related disorder, e.g. systemic lupus erythematosus,  
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.

XX  
PS Claim 1; SEQ ID NO 3369; 158bp; English.

XX  
CC The invention relates to an isolated nucleic acid encoding a PRO  
CC polypeptide. The polypeptide, agonist or an antagonist, antibody,  
CC composition, and method are useful for diagnosing and treating an immune  
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid  
CC arthritis. The present sequence represents a PRO polypeptide.

XX  
SQ Sequence 3434 BP; 1157 A; 673 C; 768 G; 836 T; 0 U; 0 Other;

Query Match 99.9%; Score 3361.8; DB 14; Length 3434;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3363; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGCGCCGCGCTGAGAGCCCTGTGACAACTCGTCATTGTTCAGGCACAGACGCGTAGAC 60  
|||||

Db 55 GCGCGCCGCGCTGAGAGCCCTGTGACAACTCGTCATTGTTCAGGCACAGACGCGTAGAC 114  
QY 61 CTGCTTCTTAAGTGGGCAGCGGACAGCCGACGCACATTTACCTGTCCCGACAGCAA 120  
|||||  
Db 115 CCTGCTTCTTAAGTGGGCAGCGGACAGCCGACGCACATTTACCTGTCCCGACAGCAA 174  
QY 121 CAGCACCATCTGCTGGGAGAACCCCTCTCCCTTCTGTAGAAAGAAAGATGTCGAATGG 180  
|||||  
Db 175 CAGCACCATCTGCTGGGAGAACCCCTCTCCCTTCTGTAGAAAGAAAGATGTCGAATGG 234  
QY 181 TATTCCACAGACGAAATTTCCGCTATCTCATCTCGTCTCAGGGCCAGGGTGAAAATG 240  
|||||  
Db 235 TATTCCACAGACGAAATTTCCGCTATCTCATCTCGTCTCAGGGCCAGGGTGAAAATG 294  
QY 241 TACATCCAGGTGAGGCTGTGCTGTGACTACCTGACCTTCTGCTGAGAGTGAAAGAG 300  
|||||  
Db 295 TACATCCAGGTGAGGCTGTGCTGTGACTACCTGACCTTCTGCTGAGAGTGAAAGAG 354  
QY 301 CAGATTCAAGAGACAGTCCGACCTCCGGGACATGACAGCAGTTGAACCTGCTGAGC 360  
|||||  
Db 355 CAGATTCAAGAGACAGTCCGACCTCCGGGACATGACAGCAGTTGAACCTGCTGAGC 414  
QY 361 ACCTTGAGAGAGGAGTCTGGCACTTGTTGACTCGGGAATTCGTGAGGCCCTCCGG 420  
|||||  
Db 415 ACCTTGAGAGAGGAGTCTGGCACTTGTTGACTCGGGAATTCGTGAGGCCCTCCGG 474  
QY 421 AGAACCCGCGAGCCCTCTGCGCCCGCCCGCTACATGAACCCGTACAGCACTTGCCCTCT 480  
|||||  
Db 475 AGAACCCGCGAGCCCTCTGCGCCCGCCCGCTACATGAACCCGTACAGCACTTGCCCTCT 534  
QY 481 CCATCGTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACCTCTTCAGCCCACT 540  
|||||  
Db 535 CCATCGTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACCTCTTCAGCCCACT 594  
QY 541 CTGCTGACAAAGCTTCTAGTTAGAGACGCTTGATTAAGTGCATGAGAGGAACCTGTTG 600  
|||||  
Db 595 CTGCTGACAAAGCTTCTAGTTAGAGACGCTTGATTAAGTGCATGAGAGGAACCTGTTG 654  
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAACAATGAAATGAATCAGGTGTA 660  
|||||  
Db 655 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAACAATGAAATGAATCAGGTGTA 714  
QY 661 AGAGAGCTACTAAAAAGGATTTGTGCAGAAAGAAAACCTGTTCTCTGCAATTTCTGAATGTT 720  
|||||  
Db 715 AGAGAGCTACTAAAAAGGATTTGTGCAGAAAGAAAACCTGTTCTCTGCAATTTCTGAATGTT 774  
QY 721 CTTCGTCAAAACAGGAAACAATGAACCTGTGCCAAGATTAAACAGGCTCTGATTGCTCAGAA 780  
|||||  
Db 775 CTTCGTCAAAACAGGAAACAATGAACCTGTGCCAAGATTAAACAGGCTCTGATTGCTCAGAA 834  
QY 781 AGCAATGCAGAGATTGGAATTTATCACAAGTTGATGTCCTCAAGTGAAGAGCAACTT 840  
|||||  
Db 835 AGCAATGCAGAGATTGGAATTTATCACAAGTTGATGTCCTCAAGTGAAGAGCAACTT 894  
QY 841 CTTTCAACCAACAGTTTACGCCAAATCTGGAGAGAGGTCTGGGGCATGAGATACTCA 900  
|||||  
Db 895 CTTTCAACCAACAGTTTACGCCAAATCTGGAGAGAGGTCTGGGGCATGAGATACTCA 954  
QY 901 TCAGAAATCATCTTTTGGCAGATTCTTCTGTAGTTTCAGAATCAGACACAAGTTGGCAGAA 960  
|||||  
Db 955 TCAGAAATCATCTTTTGGCAGATTCTTCTGTAGTTTCAGAATCAGACACAAGTTGGCAGAA 1014  
QY 961 GGAAGTGTCAAGTCTTATGATGAAAGTCTTGGACATTAACAGCAACATGGGCAGTATTCA 1020  
|||||  
Db 1015 GGAAGTGTCAAGTCTTATGATGAAAGTCTTGGACATTAACAGCAACATGGGCAGTATTCA 1074  
QY 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGAGATCCCGAGGCCA 1080  
|||||  
Db 1075 GGCACCATGGGAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGAGATCCCGAGGCCA 1134  
QY 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTGGAAAGGAAGAAAT 1140  
|||||  
Db 1135 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTGGAAAGGAAGAAAT 1194

QY	1141	ATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCAAGTGGCTGTTTACATTGCCAAG	1200
Db	1195	ATCATCATCTGCCCTCCCTACAGGGAGTGGAAAAACCAAGTGGCTGTTTACATTGCCAAG	1254
QY	1201	GATCACTTAGACAGAAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTCTTGTCAT	1260
Db	1255	GATCACTTAGACAGAAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTCTTGTCAT	1314
QY	1261	AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCATTTTGAAGAAATGG	1320
Db	1315	AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCATTTTGAAGAAATGG	1374
QY	1321	TATCGTGTATTGGATTAAAGTGGTATACCACCACTGAAAAATATCATTTCCAGAAATTGTC	1380
Db	1375	TATCGTGTATTGGATTAAAGTGGTATACCACCACTGAAAAATATCATTTCCAGAAATTGTC	1434
QY	1381	AAGTCTGTGATATTATTAATCAGTACAGCTCAAAATCCTGAAAACTCCCTCTTAACTTG	1440
Db	1435	AAGTCTGTGATATTATTAATCAGTACAGCTCAAAATCCTGAAAACTCCCTCTTAACTTG	1494
QY	1441	GAAAATGAGAGAAGATGCTGGTGTCAATTGTCAAGCTTTTCCCTCATTTATCATTTGATGAA	1500
Db	1495	GAAAATGAGAGAAGATGCTGGTGTCAATTGTCAAGCTTTTCCCTCATTTATCATTTGATGAA	1554
QY	1501	TGTCATCACACCACAAAGAAGCAGGTATTAATACATCATGAGGCATTATTTGATGCAG	1560
Db	1555	TGTCATCACACCACAAAGAAGCAGGTATTAATACATCATGAGGCATTATTTGATGCAG	1614
QY	1561	AAGTGA AAAACAATAGACTCAAGAAAAAGAAAAACAACAGTATTTCCCTCTCTCAGATA	1620
Db	1615	AAGTGA AAAACAATAGACTCAAGAAAAAGAAAAACAACAGTATTTCCCTCTCTCAGATA	1674
QY	1621	CTGGAGCTAACAGCTTCACTGCTGTTGGAGGGGCCAGAACCAAGCCAAAGCTGAAGAA	1680
Db	1675	CTGGAGCTAACAGCTTCACTGCTGTTGGAGGGGCCAGAACCAAGCTGAAGAA	1734
QY	1681	CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACTGTTAAAGAAAAAC	1740
Db	1735	CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACTGTTAAAGAAAAAC	1794
QY	1741	CTTGATCAACTGAAAAACCAATAACAGAGCCATGCAAGAAGTTGCCATTGCAGATGCA	1800
Db	1795	CTTGATCAACTGAAAAACCAATAACAGAGCCATGCAAGAAGTTGCCATTGCAGATGCA	1854
QY	1801	ACCAGAGAAGATCCATTTTAAAGAGAACTTCTAGAAATAATGACAAGATTCAAACTTAT	1860
Db	1855	ACCAGAGAAGATCCATTTTAAAGAGAACTTCTAGAAATAATGACAAGATTCAAACTTAT	1914
QY	1861	TGTCAAATGAGTCCAAATGTTCAGATTTTGGAACTCAACCCTATGAACAATGGGCCATTCAA	1920
Db	1915	TGTCAAATGAGTCCAAATGTTCAGATTTTGGAACTCAACCCTATGAACAATGGGCCATTCAA	1974
QY	1921	ATGGA AAAAAAGCTGCAAAAAAGAAATCGCAAAAGACGTGTTGTGCAGAACTTTG	1980
Db	1975	ATGGA AAAAAAGCTGCAAAAAAGAAATCGCAAAAGACGTGTTGTGCAGAACTTTG	2034
QY	1981	AGGAAGTACAATGAGGCCCTACAATTAATGACACAATTGCAATGATAGATGCGTAACT	2040
Db	2035	AGGAAGTACAATGAGGCCCTACAATTAATGACACAATTGCAATGATAGATGCGTAACT	2094
QY	2041	CATCTTGAACCTTTCTATAATGAAGAGAAGATAAGATTTCAGTCACTTATTTTGA AACAAT	2100
Db	2095	CATCTTGAACCTTTCTATAATGAAGAGAAGATAAGATTTCAGTCACTTATTAAGATGAT	2154
QY	2101	AGTGATGAGGGTGGTGTATGATGATTTGTGATGGTGTGAAGATGAGGATGATTTTAAAG	2160
Db	2155	AGTGATGAGGGTGGTGTATGATGATTTGTGATGGTGTGAAGATGAGGATGATTTTAAAG	2214
QY	2161	AAACCTTTGA AACTGGATGA AACAAGATAGATTTCTCATGACTTTATTTTGA AACAAT	2220
Db	2215	AAACCTTTGA AACTGGATGA AACAAGATAGATTTCTCATGACTTTATTTTGA AACAAT	2274

QY	2221	AAAAATTGTTGAAAAAGCTGGCTGAAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAAAATTA	2280
Db	2275	AAAAATGTTGAAAAAGCTGGCTGAAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAAAATTA	2334
QY	2281	AGAAATATCCATAATGAGCAATATACTAGACTGAGGAATCAGCAGAGAAATATCTTT	2340
Db	2335	AGAAATATCCATAATGAGCAATATACTAGACTGAGGAATCAGCAGAGAAATATCTTT	2394
QY	2341	ACAAAAACACGACAGAGTGCATATATGCGCTTTCCCACTGCAATTACTGAAAATGAAAAATTT	2400
Db	2395	ACAAAAACACGACAGAGTGCATATATGCGCTTTCCCACTGCAATTACTGAAAATGAAAAATTT	2454
QY	2401	GCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGAGCTGGACACAGCAGTGAATTCAAA	2460
Db	2455	GCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGAGCTGGACACAGCAGTGAATTCAAA	2514
QY	2461	CCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAAATTTCCGCACTGGAATAATCAAT	2520
Db	2515	CCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAAATTTCCGCACTGGAATAATCAAT	2574
QY	2521	CTGCTTATCGCTACCAAGTGGCAGAAAGGTCAGATATTAAAGAAATGTAACTTGT	2580
Db	2575	CTGCTTATCGCTACCAAGTGGCAGAAAGGTCAGATATTAAAGAAATGTAACTTGT	2634
QY	2581	ATCCGTTATGTCCTGCTCAACCAATGAATAGCCATGCTCCAGGCCCGTGTGAGCCAGA	2640
Db	2635	ATCCGTTATGTCCTGCTCAACCAATGAATAGCCATGCTCCAGGCCCGTGTGAGCCAGA	2694
QY	2641	GCTGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG	2700
Db	2695	GCTGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG	2754
QY	2701	ACAGTTAATGATTTCCGAGAGAAAGATGATGTATTAAGCTATACATTGTGTTCAAAATATG	2760
Db	2755	ACAGTTAATGATTTCCGAGAGAAAGATGATGTATTAAGCTATACATTGTGTTCAAAATATG	2814
QY	2761	AAACGAGAGAGTATGCTCATTAAGATTTTGAATTAACAGATGCCAAAGTATTAATGAAAAAG	2820
Db	2815	AAACGAGAGAGTATGCTCATTAAGATTTTGAATTAACAGATGCCAAAGTATTAATGAAAAAG	2874
QY	2821	AAAAATGAAAAACCAAGAGAAATATTGCCAAGCATTTACAAGAAATTAACCCATCACTAAATACT	2880
Db	2875	AAAAATGAAAAACCAAGAGAAATATTGCCAAGCATTTACAAGAAATTAACCCATCACTAAATACT	2934
QY	2881	TTCTTTGCAAAAACTGCAAGTGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAAATTGAG	2940
Db	2935	TTCTTTGCAAAAACTGCAAGTGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAAATTGAG	2994
QY	2941	AAAAATGCATCAGCTCAATATGACCCAGAAATTCAGGAACCTTACATTGTAAAGAAAAAC	3000
Db	2995	AAAAATGCATCAGCTCAATATGACCCAGAAATTCAGGAACCTTACATTGTAAAGAAAAAC	3054
QY	3001	AAAGCACTGCAAAAAGAAAGTGTGCCGACTATCAAAATAAATGCTGAATCATCTGCAAAATGT	3060
Db	3055	AAAGCACTGCAAAAAGAAAGTGTGCCGACTATCAAAATAAATGCTGAATCATCTGCAAAATGT	3114
QY	3061	GGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGCTGCTCAAAATA	3120
Db	3115	GGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGCTGCTCAAAATA	3174
QY	3121	AGGAATTTTGTAGTGGTTTCAAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGTA	3180
Db	3175	AGGAATTTTGTAGTGGTTTCAAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGTA	3234
QY	3181	GAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTATAGTAGAG	3240
Db	3235	GAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTATAGTAGAG	3294
QY	3241	GATTAGCACTTGATTGAAGATTTCTTTAAATATCTACAGTTAAACATTTAATATGATTA	3300
Db	3295	GATTAGCACTTGATTGAAGATTTCTTTAAATATCTACAGTTAAACATTTAATATGATTA	3354
QY	3301	TGATTAAATGATTCATTAATGCTACAGAACTGACATTAAGAAATCAATAAATGATGTTTTA	3360

Db 3355 TGATTATGTATTATTATGCTACAGAACTGACATAGAATCAATATAATGATTGTTTA 3414  
QY 3361 CTCTG 3365  
Db 3415 CTCTG 3419

RESULT 7  
ADY20376  
ID ADY20376 standard; DNA; 3434 BP.

AC ADY20376;  
XX 05-MAY-2005 (first entry)  
XX

DE DNA encoding a PRO polypeptide, SEQ ID NO 6182.

XX Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;  
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;  
KW Antichyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;  
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antilasthmatic;  
KW Antiallergic; ds; gene; diagnosis.

XX Homo sapiens.  
XX WO2005016962-A2.

PD 24-FEB-2005.

PF 11-AUG-2004; 2004WO-US026249.

PR 11-AUG-2003; 2003US-0493546P.

XX (GETH ) GENENTECH INC.

PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;

DR WPI; 2005-182330/19.

PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and  
PT treating an immune related disorder, e.g. systemic lupus erythematosus,  
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.

PS Claim 1; SEQ ID NO 6182; 158bp; English.

XX The invention relates to an isolated nucleic acid encoding a PRO  
CC polypeptide. The polypeptide, agonist or an antagonist, antibody,  
CC composition, and method are useful for diagnosing and treating an immune  
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid  
CC arthritis. The present sequence represents a PRO polypeptide.

XX Sequence 3434 BP; 1157 A; 673 C; 768 G; 836 T; 0 U; 0 Other;

QY Query Match 99.9%; Score 3361.8; DB 14; Length 3434;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3363; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGCGCCGGCCTGAGAGCCCTGTGACAACCTGTCATTGTCAGGCACAGCGGTAGAC 60  
Db 55 GCGCGCCGGCCTGAGAGCCCTGTGACAACCTGTCATTGTCAGGCACAGCGGTAGAC 114  
QY 61 CTGCTTCTTAAGTGGGCGAGCGGACAGCGGCACATTTCACCTGTCCGACAGCAA 120  
Db 115 CTGCTTCTTAAGTGGGCGAGCGGACAGCGGCACATTTCACCTGTCCGACAGCAA 174  
QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTGAGAAAGAAAGATGCGAATGG 180  
Db 175 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTGAGAAAGAAAGATGCGAATGG 234  
QY 181 TATTCCACAGACGAAATTCGGCTATCTCATCTGCTTCAAGGCGCAGGGTGAATG 240  
Db 235 TATTCCACAGACGAAATTCGGCTATCTCATCTGCTTCAAGGCGCAGGGTGAATG 294

QY 241 TACATCCAGGTGAGCCTGTGCTGAGACTAAGCTTCTGCTGACAGAGGTGAAGAG 300  
Db 295 TACATCCAGGTGAGCCTGTGCTGAGACTAAGCTTCTGCTGACAGAGGTGAAGAG 354  
QY 301 CAGATTCAAGAGCAGTCCCACTCCGGGAAACATGACAGGCAAGTTGAAGTCTGCTGAGC 360  
Db 355 CAGATTCAAGAGCAGTCCCACTCCGGGAAACATGACAGGCAAGTTGAAGTCTGCTGAGC 414  
QY 361 ACCTTGAGAAAGGAGTCTGCGACCTGTTGAGACTCGGAAATTGTTGAGGCCCTCCGG 420  
Db 415 ACCTTGAGAAAGGAGTCTGCGACCTGTTGAGACTCGGAAATTGTTGAGGCCCTCCGG 474  
QY 421 AGAACCGGACGCCCTCTGCGCGCCGCTACATGAACCTGAGCTCACGGAAGTCCCTCT 480  
Db 475 AGAACCGGACGCCCTCTGCGCGCCGCTACATGAACCTGAGCTCACGGAAGTCCCTCT 534  
QY 481 CCATCGTTGAGAACGCTCATGATGAATATCTCCAAGTGTGAACCTCCTTACGCCCACT 540  
Db 535 CCATCGTTGAGAACGCTCATGATGAATATCTCCAAGTGTGAACCTCCTTACGCCCACT 594  
QY 541 CTGCTGACAAAGCTTCTAGTTAGAGAGCTCTTGATTAAGTGAAGAGAGAACTGTTG 600  
Db 595 CTGCTGACAAAGCTTCTAGTTAGAGAGCTCTTGATTAAGTGAAGAGAGAACTGTTG 654  
QY 601 ACAATTGAAGACAGAAACCGGATGCTGTCAGAAAAACAATGAATCAGGTGA 660  
Db 655 ACAATTGAAGACAGAAACCGGATGCTGTCAGAAAAACAATGAATCAGGTGA 714  
QY 661 AGAGAGCTACTAAAAAGGATTGTGAGAAAAAGAACTGTTCTGCAATTTCTGAATGTT 720  
Db 715 AGAGAGCTACTAAAAAGGATTGTGAGAAAAAGAACTGTTCTGCAATTTCTGAATGTT 774  
QY 721 CTTGCTCAACAGGAAACAATGAACCTGTCGAAGTTAACAGGCTCTGATTGCTCAGAA 780  
Db 775 CTTGCTCAACAGGAAACAATGAACCTGTCGAAGTTAACAGGCTCTGATTGCTCAGAA 834  
QY 781 AGCAATGACAGATTGAGAAATTTATCACAAGTTGATGTCCTCAAGTGAAGACCACTT 840  
Db 835 AGCAATGACAGATTGAGAAATTTATCACAAGTTGATGTCCTCAAGTGAAGACCACTT 894  
QY 841 CTTTCAACCAAGTTTCAAGCCAAATCTGAGAAAGAGGCTTGCGGCATGAGAAATACTCA 900  
Db 895 CTTTCAACCAAGTTTCAAGCCAAATCTGAGAAAGAGGCTTGCGGCATGAGAAATACTCA 954  
QY 901 TCAGAAATCATCTTTGACAGATTCTTCTGATTTCAGAAATCAGACACAAGTTGGCAGAA 960  
Db 955 TCAGAAATCATCTTTGACAGATTCTTCTGATTTCAGAAATCAGACACAAGTTGGCAGAA 1014  
QY 961 GGAAGTGCAGCTGCTTAGATGAAAGTCTTGGACATTAACAGACAAATGGGAGTATTC 1020  
Db 1015 GGAAGTGCAGCTGCTTAGATGAAAGTCTTGGACATTAACAGACAAATGGGAGTATTC 1074  
QY 1021 GGCAACCATGGGAAGTATTGAGATGAAGAAATGTGGCAGCAAGCATCCCGGAGCCA 1080  
Db 1075 GGCAACCATGGGAAGTATTGAGATGAAGAAATGTGGCAGCAAGCATCCCGGAGCCA 1134  
QY 1081 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCCTTGAAGGGAAGAA 1140  
Db 1135 GAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCCTTGAAGGGAAGAA 1194  
QY 1141 ATCATCATCTGCTCCCTACAGGAGTGAAGAAAAACAGAGTGGCTTTACATTGGCAAG 1200  
Db 1195 ATCATCATCTGCTCCCTACAGGAGTGAAGAAAAACAGAGTGGCTTTACATTGGCAAG 1254  
QY 1201 GATCACTTAGACAGAGAAAGAAAGCATCTGAGCCTGGAAGATTTTGAAGAAATG 1260  
Db 1255 GATCACTTAGACAGAGAAAGAAAGCATCTGAGCCTGGAAGATTTTGAAGAAATG 1314  
QY 1261 AAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCACCAACATTTTGAAGAAATG 1320  
Db 1315 AAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCACCAACATTTTGAAGAAATG 1374

QY 1321 TATCGTGTATTGATTAAGTGTGATACCACCACTGAAATATCATTTCCAGAAGTTGTC 1380  
|||||  
Db 1375 TATCGTGTATTGATTAAGTGTGATACCACCACTGAAATATCATTTCCAGAAGTTGTC 1434  
QY 1381 AAGTCTGTGATATTATATCATGATACAGCTCAAAATCCTTGAAAACCTCCCTCTTAAACTTG 1440  
|||||  
Db 1435 AAGTCTGTGATATTATATCATGATACAGCTCAAAATCCTTGAAAACCTCCCTCTTAAACTTG 1494  
QY 1441 GAAAATGAGAGAAGATGCTGTGTCAATTGTTCAGACTTTCCCTCATTTATCATTTGATGAA 1500  
|||||  
Db 1495 GAAAATGAGAGAAGATGCTGTGTCAATTGTTCAGACTTTCCCTCATTTATCATTTGATGAA 1554  
QY 1501 TGTTCATCACACCAACAAGAGAGAGTGTATTAATACATCATGAGGCATTATTGATGCAG 1560  
|||||  
Db 1555 TGTTCATCACACCAACAAGAGAGAGTGTATTAATACATCATGAGGCATTATTGATGCAG 1614  
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAACCAACAGTGATTTCCCTTCTCAGATA 1620  
|||||  
Db 1615 AAGTTGAAAAACAATAGACTCAAGAAAGAAAACCAACAGTGATTTCCCTTCTCAGATA 1674  
QY 1621 CTGGGACTAACAGCTTCACTGTGTGTGGAGGGGCCAGAAAGCAAGCCAAAGCTGAAGAA 1680  
|||||  
Db 1675 CTGGGACTAACAGCTTCACTGTGTGTGGAGGGGCCAGAAAGCAAGCCAAAGCTGAAGAA 1734  
QY 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTACTATTAAAACTGTGTTAAAGAAAAC 1740  
|||||  
Db 1735 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTACTATTAAAACTGTGTTAAAGAAAAC 1794  
QY 1741 CTTCATCAACTGAAAAACCAAAATACAGAGCCATGCAAGAAGTTGCCATTGCAGATGCA 1800  
|||||  
Db 1795 CTTCATCAACTGAAAAACCAAAATACAGAGCCATGCAAGAAGTTGCCATTGCAGATGCA 1854  
QY 1801 ACCAGAGAAGATCCATTTTAAAGAGAAAACCTTCTAGAAATATGACAGAAGATTCAAACCTTAT 1860  
|||||  
Db 1855 ACCAGAGAAGATCCATTTTAAAGAGAAAACCTTCTAGAAATATGACAGAAGATTCAAACCTTAT 1914  
QY 1861 TGTCAAAATGAGTCCAAATGTCAAGATTTTGGAACTCAACCTATGAACAATGGGCCATTCAA 1920  
|||||  
Db 1915 TGTCAAAATGAGTCCAAATGTCAAGATTTTGGAACTCAACCTATGAACAATGGGCCATTCAA 1974  
QY 1921 ATGCAAAAAAAGCTGCAAAAAAAGAAATCGCAAGAAAGCTGTTGTGCAAAACATTTG 1980  
|||||  
Db 1975 ATGCAAAAAAAGCTGCAAAAAAAGAAATCGCAAGAAAGCTGTTGTGCAAAACATTTG 2034  
QY 1981 AGGAAGTACAATGAGGCCCTCAAAATTAATGACACAATTGCAATGATAGATGCGTATACT 2040  
|||||  
Db 2035 AGGAAGTACAATGAGGCCCTCAAAATTAATGACACAATTGCAATGATAGATGCGTATACT 2094  
QY 2041 CATCTTGAACCTTCTATATGAAGAGAAAGATAGAAGTTTGCACTCATAGAAAGATGAT 2100  
|||||  
Db 2095 CATCTTGAACCTTCTATATGAAGAGAAAGATAGAAGTTTGCACTCATAGAAAGATGAT 2154  
QY 2101 AGTGATGAGGGTGTGATGATGATGATTTGTGATGGTGAAGATGAGATGATTTAAAG 2160  
|||||  
Db 2155 AGTGATGAGGGTGTGATGATGATGATTTGTGATGGTGAAGATGAGATGATTTAAAG 2214  
QY 2161 AAACCTTTGAACCTGATGAACAGATAGATTTCTCATGACTTTATTTTGA AAAACAAT 2220  
|||||  
Db 2215 AAACCTTTGAACCTGATGAACAGATAGATTTCTCATGACTTTATTTTGA AAAACAAT 2274  
QY 2221 AAAATGTTGAAAAGCTGGCTGAAAACCCAGAATATGAAAATGAAAAGCTGACCAAAATTA 2280  
|||||  
Db 2275 AAAATGTTGAAAAGCTGGCTGAAAACCCAGAATATGAAAATGAAAAGCTGACCAAAATTA 2334  
QY 2281 AGAATAATCCATAATGAGACATATACTAGAGCTGAGGAATCAGCAGAGAAATAATCTTT 2340  
|||||  
Db 2335 AGAATAATCCATAATGAGACATATACTAGAGCTGAGGAATCAGCAGAGAAATAATCTTT 2394  
QY 2341 ACAAAAAACAGCAGAGATGATATGCGCTTTCCAGTGGAATTACTGAAAAATGAAAAATTT 2400  
|||||  
Db 2395 ACAAAAAACAGCAGAGATGATATGCGCTTTCCAGTGGAATTACTGAAAAATGAAAAATTT 2454  
QY 2401 GCTGAAGTAGAGTCAAAAGCCCAACATCTGATTGAGCTGGAACAGCAGAGTGAGTTCAAA 2460

Db 2455 GCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGGAGCTGACACAGCAGTGAGTTCAAA 2514  
QY 2461 CCCATGACACAGAATGAACAAAAGAGTCAATTAGTAATTTGCACTGAAAAATCAAT 2520  
|||||  
Db 2515 CCCATGACACAGAATGAACAAAAGAGTCAATTAGTAATTTGCACTGAAAAATCAAT 2574  
QY 2521 CTGCTTATCGCTACCAAGTGGCAGAGAAGGCTGATATATTAAAGATGTAACATTGTT 2580  
|||||  
Db 2575 CTGCTTATCGCTACCAAGTGGCAGAGAAGGCTGATATATTAAAGATGTAACATTGTT 2634  
QY 2581 ATCCGTATGTGCTCGTCAACCAATGAATAGCCATGCTCCAGGCCCGTGTGAGCCAGA 2640  
|||||  
Db 2635 ATCCGTATGTGCTCGTCAACCAATGAATAGCCATGCTCCAGGCCCGTGTGAGCCAGA 2694  
QY 2641 GCTGATGAGAGCACTACGTCCTGTGTGCTCAAGTGTTCAAGAGTTATCGAACATGAG 2700  
|||||  
Db 2695 GCTGATGAGAGCACTACGTCCTGTGTGCTCAAGTGTTCAAGAGTTATCGAACATGAG 2754  
QY 2701 ACAGTTAATGATTTCCGAGAGAGATGATGATTAAGCTATACATTTGTCAAAATATG 2760  
|||||  
Db 2755 ACAGTTAATGATTTCCGAGAGAGATGATGATTAAGCTATACATTTGTCAAAATATG 2814  
QY 2761 AAACCAAGAGAGTATGCTCATAGATTTTGAATTAACAGATGCAAGATGTAATGAAAAAG 2820  
|||||  
Db 2815 AAACCAAGAGAGTATGCTCATAGATTTTGAATTAACAGATGCAAGATGTAATGAAAAAG 2874  
QY 2821 AAAATGA AAAACCAAGAGAAATATTGCCAAGCAATTACAAGATTAACCACTAATAACT 2880  
|||||  
Db 2875 AAAATGA AAAACCAAGAGAAATATTGCCAAGCAATTACAAGATTAACCACTAATAACT 2934  
QY 2881 TTCTTTGCAAAAACTGCAAGTGTGTAGCTGTCTGTGGGAAGATATCATGTAATTGAG 2940  
|||||  
Db 2935 TTCTTTGCAAAAACTGCAAGTGTGTAGCTGTCTGTGGGAAGATATCATGTAATTGAG 2994  
QY 2941 AAAATGCATCAGTCAATATGACCCCAAGAAATCAAGAACTTTACATTGTAAGAGAAAAC 3000  
|||||  
Db 2995 AAAATGCATCAGTCAATATGACCCCAAGAAATCAAGAACTTTACATTGTAAGAGAAAAC 3054  
QY 3001 AAAGCACTGC AAAAGAGTGTGCCGACTATCAATAATGTTGAATCATCTGCAATGT 3060  
|||||  
Db 3055 AAAGCACTGC AAAAGAGTGTGCCGACTATCAATAATGTTGAATCATCTGCAATGT 3114  
QY 3061 GGCCAGGCTTGGGGAACAATGATGTGCAAAAGGCTTAGATTTGCTGTCTCAAAATA 3120  
|||||  
Db 3115 GGCCAGGCTTGGGGAACAATGATGTGCAAAAGGCTTAGATTTGCTGTCTCAAAATA 3174  
QY 3121 AGGAATTTTGTAGTGTTC AAAAATTAATCAACAAAGAAACAATACAAAAAGTGCGTA 3180  
|||||  
Db 3175 AGGAATTTTGTAGTGTTC AAAAATTAATCAACAAAGAAACAATACAAAAAGTGCGTA 3234  
QY 3181 GAATTAACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTTATTTAGTATGAG 3240  
|||||  
Db 3235 GAATTAACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTTATTTAGTATGAG 3294  
QY 3241 GATTAGCACTTGATTGAAGATTTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300  
|||||  
Db 3295 GATTAGCACTTGATTGAAGATTTCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3354  
QY 3301 TGATTAATGTATTCAATTAATGCTACAGAACTGCATTAAGAAATCAATTAATAATGATTGTTTA 3360  
|||||  
Db 3355 TGATTAATGTATTCAATTAATGCTACAGAACTGCATTAAGAAATCAATTAATAATGATTGTTTA 3414  
QY 3361 CTCTG 3365  
|||||  
Db 3415 CTCTG 3419

RESULT 8  
ACN91892  
ID ACN91892 standard; DNA; 3928 BP.  
XX  
AC ACN91892;

XX 02-DEC-2004 (first entry)  
XX Breast cancer related marker, seq id 13042.  
XX Cancer; breast; tumour; cytoslastic; marker; detection; therapy; ds.  
XX Homo sapiens.  
XX US2003099974-A1.  
XX 29-MAY-2003.  
XX 18-JUL-2002; 2002US-00198846.  
XX 18-JUL-2001; 2001US-0306220P.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX Lillie J, Xu Y, Wang Y, Steinmann K;  
XX WPI; 2003-787014/74.  
XX Novel isolated polypeptide associated with breast cancer, useful for  
PT detecting presence of polypeptide in sample, as a marker for breast  
PT cancer.  
XX  
XX Disclosure; SEQ ID NO 13042; 36pp; English.  
XX  
XX The invention relates to an isolated polypeptide (I) associated with  
CC breast cancer which is encoded by a nucleic acid molecule comprising a  
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to  
CC the polypeptide of the invention. The activity of the polypeptide of the  
CC invention may be described as cytoslastic. The antibody is useful for  
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the  
CC invention are useful in the detection of breast tumours. (I) is useful as  
CC a marker for breast cancer and in breast cancer therapy. Sequences given  
CC in records ACN78851-ACN92934 represent nucleic acid markers associated  
CC with breast cancer. Note: The sequence listing does not form part of the  
CC specification but may be obtained in electronic format from the USPTO web  
CC site at [Seqdata.uspto.gov/sequence.html?docID=20030099974](http://Seqdata.uspto.gov/sequence.html?docID=20030099974)  
XX  
SQ Sequence 3928 BP; 1285 A; 826 C; 895 G; 917 T; 0 U; 5 Other;  
  
Query Match 99.9%; Score 3360.2; DB 11; Length 3928;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3362; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 GCGGCCGCGCTGAGAGCCCTGTGACAACCTGTCATGTGTGAGGACAGAGCGGTAGAC 60  
Db 443 GCGGCCGCGCTGAGAGCCCTGTGACAACCTGTCATGTGTGAGGACAGAGCGGTAGAC 502  
QY 61 CCTGCTTCTTAAGTGGGACGGGACAGCGGACGCAATTTCACTGTCCCGCAGACAA 120  
Db 503 CCTGCTTCTTAAGTGGGACGGGACAGCGGACGCAATTTCACTGTCCCGCAGACAA 562  
QY 121 CAGCACCATCTGCTTGGGAGAACTCTCCCTTCTGTGAGAAAGAAAGATGTGAATGGG 180  
Db 563 CAGCACCATCTGCTTGGGAGAACTCTCCCTTCTGTGAGAAAGAAAGATGTGAATGGG 622  
QY 181 TATTCACAGACGAGAAATTCGCTATCTCATCTGCTCTCAGGGCCAGGCTGAAAATG 240  
Db 623 TATTCACAGACGAGAAATTCGCTATCTCATCTGCTCTCAGGGCCAGGCTGAAAATG 682  
QY 241 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCTGCTGACAGGTGAAGAG 300  
Db 683 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCTGCTGACAGGTGAAGAG 742  
QY 301 CAGATTCAAGACAGCAGTCCGCACTCCGGGAACATGACGCAAGTGAAGTGTGCTGAGC 360  
Db 743 CAGATTCAAGACAGCAGTCCGCACTCCGGGAACATGACGCAAGTGAAGTGTGCTGAGC 802  
QY 361 ACCTTGAGAGAGGAGTGTGGCACTTGTGTGACTGGGGAATTCGTGAGGCCCTCCGG 420

Db 803 ACCTTGAGAGAGGAGTGTGGCACTTGTGTGACTCGGGAATTCGTGAGGCCCTCCGG 862  
QY 421 AGAACCGGACGCCCTCTGCGCCGCCGCTACATGAACCTGAGCTCAGCACTTGCCCTCT 480  
Db 863 AGAACCGGACGCCCTCTGCGCCGCCGCTACATGAACCTGAGCTCAGCACTTGCCCTCT 922  
QY 481 CCATCGTTTGAGAACGCTCATGATGAAATATCTCCAAGCTGTGAACCTCTCAGCCCACT 540  
Db 923 CCATCGTTTGAGAACGCTCATGATGAAATATCTCCAAGCTGTGAACCTCTCAGCCCACT 982  
QY 541 CTGCTGACAAAGCTTCTAGTTAGAGACGCTTGTGATTAAGTGCATGAGAGGAAGTGTG 600  
Db 983 CTGCTGACAAAGCTTCTAGTTAGAGACGCTTGTGATTAAGTGCATGAGAGGAAGTGTG 1042  
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGTGAGAAACCAATGAAATGAATCAGTGTGA 660  
Db 1043 ACAATTGAAGACAGAAACCGGATTGCTGTGAGAAACCAATGAAATGAATCAGTGTGA 1102  
QY 661 AGAGAGCTACTAAAAAGGATTGTGACAGAAAGAAACTGCTTCTGCAATTTCTGAATGTT 720  
Db 1103 AGAGAGCTACTAAAAAGGATTGTGACAGAAAGAAACTGCTTCTGCAATTTCTGAATGTT 1162  
QY 721 CTTGCTCAAAACAGAAACAATGAACCTGTGCAAGAGTTAACAGGCTCTGATTGCTCAGAA 780  
Db 1163 CTTGCTCAAAACAGAAACAATGAACCTGTGCAAGAGTTAACAGGCTCTGATTGCTCAGAA 1222  
QY 781 AGCAATGACAGATGAGAAATTTATCAACAAGTTGATGTTCTCAAGTGAAGACCACTT 840  
Db 1223 AGCAATGACAGATGAGAAATTTATCAACAAGTTGATGTTCTCAAGTGAAGACCACTT 1282  
QY 841 CTTTCAACCAAGTTTCAAGCCCAATCTGAGAGAGAGTCTGGGGCATGAGATACTCA 900  
Db 1283 CTTTCAACCAAGTTTCAAGCCCAATCTGAGAGAGAGTCTGGGGCATGAGATACTCA 1342  
QY 901 TCAGAAATCATCTTTTGAGATTTCTTGTGATTTTCAGAATCAGACACAAAGTTGGCAGAA 960  
Db 1343 TCAGAAATCATCTTTTGAGATTTCTTGTGATTTTCAGAATCAGACACAAAGTTGGCAGAA 1402  
QY 961 GGAAGTGTGAGCTGCTTGAATGAAGTCTTGACATTAACAGCAACATGGGCAAGTATCA 1020  
Db 1403 GGAAGTGTGAGCTGCTTGAATGAAGTCTTGACATTAACAGCAACATGGGCAAGTATCA 1462  
QY 1021 GGCACCATGGAAGTATTCAGATGAAGAGATGTGGCAGCAAGACATCCCGGAGCCA 1080  
Db 1463 GGCACCATGGAAGTATTCAGATGAAGAGATGTGGCAGCAAGACATCCCGGAGCCA 1522  
QY 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCTTGAAGGGAAGAAAT 1140  
Db 1523 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCTTGAAGGGAAGAAAT 1582  
QY 1141 ATCATCATCTGCTCCCTTACAGGGAGTGAAGAAACAGAGTGGCTTTTACATTGCCAAG 1200  
Db 1583 ATCATCATCTGCTCCCTTACAGGGAGTGAAGAAACAGAGTGGCTTTTACATTGCCAAG 1642  
QY 1201 GATCACTTAGACAAGAAAGAAAGCATCTGAGCCTGGAAGAGTTATAGTTCTGTCAAT 1260  
Db 1643 GATCACTTAGACAAGAAAGAAAGCATCTGAGCCTGGAAGAGTTATAGTTCTGTCAAT 1702  
QY 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAAACATTTTGAAGAAATGG 1320  
Db 1703 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAAACATTTTGAAGAAATGG 1762  
QY 1321 TATCGTGTATTTGATTAAGTGTGATACCCAATGAAATATATCTTCCAGAGTGTGC 1380  
Db 1763 TATCGTGTATTTGATTAAGTGTGATACCCAATGAAATATATCTTCCAGAGTGTGC 1822  
QY 1381 AAGTCTGTGATATTTATATCAGTACAGCTCAAAATCCTTGAAACATCTCTTAAACTTG 1440  
Db 1823 AAGTCTGTGATATTTATATCAGTACAGCTCAAAATCCTTGAAACATCTCTTAAACTTG 1882  
QY 1441 GAAATGAGAGAGATGTGTGTTCAATGTGCACTTTTCCCTCATATCATGTGATGAA 1500

Db 1883 GAAATGAGAGATGCTGCTGTTCAATTGTCAAGACTTTTCCCTCATTTATCATGTGAA 1942  
Qy 1501 TGTGATCACACCAACAAGAGCAGTGTATATATACATCATGAGGCATTATTTGATGAG 1560  
Db 1943 TGTTCATCACACCAACAAGAGCAGTGTATATATACATCATGAGGCATTATTTGATGAG 2002  
Qy 1561 AAGTTGAAAAACAATAGACTCAAGAAAAACAACCAAGTATTCCTCTCAGATA 1620  
Db 2003 AAGTTGAAAAACAATAGACTCAAGAAAAACAACCAAGTATTCCTCTCAGATA 2062  
Qy 1621 CTGGGACTAACAGCTTCACTGCTGTTGGAGGGGCCAAGCAAGCCAAAGCTGAAGAA 1680  
Db 2063 CTGGGACTAACAGCTTCACTGCTGTTGGAGGGGCCAAGCAAGCCAAAGCTGAAGAA 2122  
Qy 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACTGTTAAAGAAAC 1740  
Db 2123 CACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACTGTTAAAGAAAC 2182  
Qy 1741 CTTGATCAACTGAAAAACCAATACAGAGCCATGCAAGAGTTGCCATTGCAGATGCA 1800  
Db 2183 CTTGATCAACTGAAAAACCAATACAGAGCCATGCAAGAGTTGCCATTGCAGATGCA 2242  
Qy 1801 ACCAGAGAAGATCCATTTAAAGAGAACTTCTAGAAATATGACAAGGATTCAAACTTAT 1860  
Db 2243 ACCAGAGAAGATCCATTTAAAGAGAACTTCTAGAAATATGACAAGGATTCAAACTTAT 2302  
Qy 1861 TGTCAATGAGTCCAAATGTGAGATTTTGGAACTCAACCTATGAACAATGGCCATTCAA 1920  
Db 2303 TGTCAATGAGTCCAAATGTGAGATTTTGGAACTCAACCTATGAACAATGGCCATTCAA 2362  
Qy 1921 ATGGAATAAAAGCTGCATAAAAGGAATCGCAAGAGACGTGTTGTGCAACAATTG 1980  
Db 2363 ATGGAATAAAAGCTGCATAAAAGGAATCGCAAGAGACGTGTTGTGCAACAATTG 2422  
Qy 1981 AGGAAGTACAATGAGGCGCTTACAATTATGACACAACTTGAATGATAGTCGTATACT 2040  
Db 2423 AGGAAGTACAATGAGGCGCTTACAATTATGACACAACTTGAATGATAGTCGTATACT 2482  
Qy 2041 CATCTGAAAACCTTCTATTAATGAAGAGAAAGATTAAGATTGCACTCATAGAAGATGAT 2100  
Db 2483 CATCTGAAAACCTTCTATTAATGAAGAGAAAGATTAAGATTGCACTCATAGAAGATGAT 2542  
Qy 2101 AGTGATGAGGGTGTGATGATGATATTGTGATGGTGAAGATGAGGATGATTAAAG 2160  
Db 2543 AGTGATGAGGGTGTGATGATGATATTGTGATGGTGAAGATGAGGATGATTAAAG 2602  
Qy 2161 AAACCTTTGAAACTGGATGAACAAGATGATTTCTCATGACTTTATTTTGAACAAT 2220  
Db 2603 AAACCTTTGAAACTGGATGAACAAGATGATTTCTCATGACTTTATTTTGAACAAT 2662  
Qy 2221 AAAATGTTGAAAAAGCTGCTGTA AAAACCCAGAATATGAAAAAGCTGACCAATTA 2280  
Db 2663 AAAATGTTGAAAAAGCTGCTGTA AAAACCCAGAATATGAAAAAGCTGACCAATTA 2722  
Qy 2281 AGAATAACCATATATGAGCAATATACTAGACTGAGGAATCAGCAGGAAATTAATCTTT 2340  
Db 2723 AGAATAACCATATATGAGCAATATACTAGACTGAGGAATCAGCAGGAAATTAATCTTT 2782  
Qy 2341 ACAAATAACAGACAGAGTGCATATGCGCTTTCCCACTGGAATTACTGAAAAATGA AAAATT 2400  
Db 2783 ACAAATAACAGACAGAGTGCATATGCGCTTTCCCACTGGAATTACTGAAAAATGA AAAATT 2842  
Qy 2401 GCTGAAGTAGAGTCAAAAGCCCACTCTGATTGAGCTGGAACACAGCAGTGAAGTCAAA 2460  
Db 2843 GCTGAAGTAGAGTCAAAAGCCCACTCTGATTGAGCTGGAACACAGCAGTGAAGTCAAA 2902  
Qy 2461 CCCATGACACAGAATGAACAAAAAGAGTCAATTAGTAATTTGCACTGAAAAATCAAT 2520  
Db 2903 CCCATGACACAGAATGAACAAAAAGAGTCAATTAGTAATTTGCACTGAAAAATCAAT 2962  
Qy 2521 CTGCTTATCGCTACACAGTGGCAGAGAAGAGTCTGATATTTAAAGAATGTAACTTTGTT 2580  
Db 2963 CTGCTTATCGCTACACAGTGGCAGAGAAGAGTCTGATATTTAAAGAATGTAACTTTGTT 3022

Qy 2581 ATCCGTTATGCTCTCGTCAACCAATGAATAATAGCCATGCTCCAGGCCCGTGTGAGCCAGA 2640  
Db 3023 ATCCGTTATGCTCTCGTCAACCAATGAATAATAGCCATGCTCCAGGCCCGTGTGAGCCAGA 3082  
Qy 2641 GCTGATGAGAGCACCTACGTCCTGTTGCTGCTCAGAGTGTTCAGAGTTATCGAACATGAG 2700  
Db 3083 GCTGATGAGAGCACCTACGTCCTGTTGCTGCTCAGAGTGTTCAGAGTTATCGAACATGAG 3142  
Qy 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGTATTAAGCTATACATTTGTTC AAAATATG 2760  
Db 3143 ACAGTTAATGATTTCCGAGAGAAGATGATGTATTAAGCTATACATTTGTTC AAAATATG 3202  
Qy 2761 AAACGAGAGAGTATGCTCATTAAGATTTTGGAAATTAACAGATGCAAAAGTATTAATGAAAAAG 2820  
Db 3203 AAACGAGAGAGTATGCTCATTAAGATTTTGGAAATTAACAGATGCAAAAGTATTAATGAAAAAG 3262  
Qy 2821 AAATGAAAAACCAAGAGAAATATTTGCCAAGCAATTAACAAGATTAACCACTAATAACT 2880  
Db 3263 AAATGAAAAACCAAGAGAAATATTTGCCAAGCAATTAACAAGATTAACCACTAATAACT 3322  
Qy 2881 TTCCTTTGCAAAAACTGAGTGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940  
Db 3323 TTCCTTTGCAAAAACTGAGTGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAATTGAG 3382  
Qy 2941 AAATGCAATCAGCTCAATATGAACCCCAAGATTCAAGGAACCTTACATTGTAAGAGAAAAAC 3000  
Db 3383 AAATGCAATCAGCTCAATATGAACCCCAAGATTCAAGGAACCTTACATTGTAAGAGAAAAAC 3442  
Qy 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAATAATTAATGTTGTAATCATCTGCCAAATGT 3060  
Db 3443 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAATAATTAATGTTGTAATCATCTGCCAAATGT 3502  
Qy 3061 GGCCAGGCTTGGGGAACAATGATGTGTCACAAAGCCTTAGATTGCTGCTCAAAATA 3120  
Db 3503 GGCCAGGCTTGGGGAACAATGATGTGTCACAAAGCCTTAGATTGCTGCTCAAAATA 3562  
Qy 3121 AGGAATTTTGTAGTGTTC AAAAATTAATTCACAAAGAACAAATACAAAAAGTGCGTA 3180  
Db 3563 AGGAATTTTGTAGTGTTC AAAAATTAATTCACAAAGAACAAATACAAAAAGTGCGTA 3622  
Qy 3181 GAATTAACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTTAGTGATGAG 3240  
Db 3623 GAATTAACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTTAGTGATGAG 3682  
Qy 3241 GATTAGCACTGATTGAAGATTCCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300  
Db 3683 GATTAGCACTGATTGAAGATTCCTTTAAATACTATCAGTTAAACATTTAATATGATTA 3742  
Qy 3301 TGAATTAATGATTCATTAATGCTACAGAACTGACATAAAGATCAATCAATTAATGATTTTA 3360  
Db 3743 TGAATTAATGATTCATTAATGCTACAGAACTGACATAAAGATCAATCAATTAATGATTTTA 3802  
Qy 3361 CTCTG 3365  
Db 3803 CTCTG 3807

RESULT 9  
ADN04879  
ID ADN04879 standard; cDNA; 3379 BP.  
XX  
AC ADN04879;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE Antipsoriatic cDNA sequence #654.  
XX  
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.  
OS Homo sapiens.  
XX  
PN WO2004028479-A2.

XX 08-APR-2004.  
PD  
XX 25-SEP-2003; 2003WO-US030907.  
PF  
XX 25-SEP-2002; 2002US-0414006P.  
PR  
XX (GETH ) GENENTECH INC.  
PA  
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI,  
PI Wu TD;  
XX  
XX WPI; 2004-305105/28.  
DR P-PSDB; ADN04880.  
DR  
XX  
PT New PRO nucleic acid or polypeptide, useful for preparing a  
PT pharmaceutical composition for diagnosing or treating psoriasis in a  
PT mammal.  
XX  
PS Claim 1; SEQ ID NO 1273; 3069pp; English.  
XX  
CC The invention relates to novel polynucleotide and polypeptides for  
CC treating psoriasis or a sequence having at least 80% identity to the  
CC above sequences. The nucleic acid is useful for preparing a composition  
CC for diagnosing or treating psoriasis in a mammal. This sequence  
CC corresponds to one of the polynucleotides of the invention.  
XX  
SQ Sequence 3379 BP; 1152 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Query Match 99.6%; Score 3353; DB 12; Length 3379;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3364; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GCGCGCGGCTGAGAGCCCTGTGGACAACCTCGTCATTTGTCAGGCACAGAGCGGTAGAC 60  
Db 1 GCGCGCGGCTGAGAGCCCTGTGGACAACCTCGTCATTTGTCAGGCACAGAGCGGTAGAC 60  
QY 61 CCTGCTTCTTAAGTGGGACGGGACAGCGGACGACATTTCACTGTCCGCGACAGCAA 120  
Db 61 CCTGCTTCTTAAGTGGGACGGGACAGCGGACGACATTTCACTGTCCGCGACAGCAA 120  
QY 121 CAGCACCATCTGCTTGGAGAACCTCTCCCTTCTGTGAGAAAGAAAGATGCGAATGG 180  
Db 121 CAGCACCATCTGCTTGGAGAACCTCTCCCTTCTGTGAGAAAGAAAGATGCGAATGG 180  
QY 181 TATTCACAGACGAGATTTCCGCTATCTCATCTCGTCTTCAAGGCCAGGGTGAATG 240  
Db 181 TATTCACAGACGAGATTTCCGCTATCTCATCTCGTCTTCAAGGCCAGGGTGAATG 240  
QY 241 TACATTCAGGTGAGAGCTGTGCTGGACTACCTGACCTTTCTGCTGACAGAGTGAAGAG 300  
Db 241 TACATTCAGGTGAGAGCTGTGCTGGACTACCTGACCTTTCTGCTGACAGAGTGAAGAG 300  
QY 301 CAGATTCAGAGGACAGTCCCACTCCGGGAACATGACGGCAGTTGAACCTGCTGAGC 360  
Db 301 CAGATTCAGAGGACAGTCCCACTCCGGGAACATGACGGCAGTTGAACCTGCTGAGC 360  
QY 361 ACCTTGAGAGAGGAGTCTGGCACTTGGTGAAGTGGGAATTCGTGAGAGCGCCCTCCGG 420  
Db 361 ACCTTGAGAGAGGAGTCTGGCACTTGGTGAAGTGGGAATTCGTGAGAGCGCCCTCCGG 420  
QY 421 AGAACCGGACGCCCTTGCGCGCCGCTACATGAACCTGAGCTCACGGACTTGCCCTCT 480  
Db 421 AGAACCGGACGCCCTTGCGCGCCGCTACATGAACCTGAGCTCACGGACTTGCCCTCT 480  
QY 481 CCATCGTTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACCTCTTCAGCCCACT 540  
Db 481 CCATCGTTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACCTCTTCAGCCCACT 540  
QY 541 CTGGTGACAAGCTTCTAGTTAGAGACGCTTGGATAGTGCATGAGAGGAACTGTTG 600  
Db 541 CTGGTGACAAGCTTCTAGTTAGAGACGCTTGGATAGTGCATGAGAGGAACTGTTG 600

QY 601 ACAATTGAGACAGAAACCGGATTGCTGTCAGAAAAACAATGAAATGAATCAGGTGA 660  
Db 601 ACAATTGAGACAGAAACCGGATTGCTGTCAGAAAAACAATGAAATGAATCAGGTGA 660  
QY 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAAGTGTCTCTGCAATTTCTGAATGT 720  
Db 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAAGTGTCTCTGCAATTTCTGAATGT 720  
QY 721 CTTGCTCAACAGGAAACAATGAACCTGTCCAGAAGTTAACAGGCTCTGATTGCTCAGAA 780  
Db 721 CTTGCTCAACAGGAAACAATGAACCTGTCCAGAAGTTAACAGGCTCTGATTGCTCAGAA 780  
QY 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGTCCTCAAGTGAAGAGCAACTT 840  
Db 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGTCCTCAAGTGAAGAGCAACTT 840  
QY 841 CTTTCAACCAAGTTCAAGCCAAATCTGAGAGAGAGGTCTGGGGCATGGAGATACTCA 900  
Db 841 CTTTCAACCAAGTTCAAGCCAAATCTGAGAGAGAGGTCTGGGGCATGGAGATACTCA 900  
QY 901 TCAGAAATCATCTTTGAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTGGCAGAA 960  
Db 901 TCAGAAATCATCTTTGAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTGGCAGAA 960  
QY 961 GGAAGTGTCAAGTCTTAGATGAAGTCTTGAACATTAACAGCAACATGGGCAGTATTC 1020  
Db 961 GGAAGTGTCAAGTCTTAGATGAAGTCTTGAACATTAACAGCAACATGGGCAGTATTC 1020  
QY 1021 GGCACCATGGGAAGTATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA 1080  
Db 1021 GGCACCATGGGAAGTATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA 1080  
QY 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTGGAAAGGGAAGAT 1140  
Db 1081 GAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTGGAAAGGGAAGAT 1140  
QY 1141 ATCATCATCTGCTCCCTACAGGAGTGAAGAAAACAGAGTGGCTTTTACATTGGCCAAG 1200  
Db 1141 ATCATCATCTGCTCCCTACAGGAGTGAAGAAAACAGAGTGGCTTTTACATTGGCCAAG 1200  
QY 1201 GATCACTTAGACAAGAAAAAAGCATCTGAGCCTTGAAAAATATATTCTTGTCAAT 1260  
Db 1201 GATCACTTAGACAAGAAAAAAGCATCTGAGCCTTGAAAAATATATTCTTGTCAAT 1260  
QY 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGATTCCAAACCATTTTGAAGAAATGG 1320  
Db 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGATTCCAAACCATTTTGAAGAAATGG 1320  
QY 1321 TATCGTGTATTGGATTAAAGTGTGATACCAACTGAAAAATATATTCTTCCAGAAATTGTC 1380  
Db 1321 TATCGTGTATTGGATTAAAGTGTGATACCAACTGAAAAATATATTCTTCCAGAAATTGTC 1380  
QY 1381 AAGTCTGTGATATTATATCAGTACAGCTCAAACTCTGAAAACTCCCTTAAACTTG 1440  
Db 1381 AAGTCTGTGATATTATATCAGTACAGCTCAAACTCTGAAAACTCCCTTAAACTTG 1440  
QY 1441 GAAAAATGAGAAAGATGCTGGTGTTCATTGTTCAGACTTTTCCCTCATTAATTGATGAA 1500  
Db 1441 GAAAAATGAGAAAGATGCTGGTGTTCATTGTTCAGACTTTTCCCTCATTAATTGATGAA 1500  
QY 1501 TGTTCATCACACCAAC -AAGAAAGAGTGTATTAATACATCATGAGGCAATTAATTGATGAG 1560  
Db 1501 TGTTCATCACACCAAC -AAGAAAGAGTGTATTAATACATCATGAGGCAATTAATTGATGAG 1560  
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAACAAAACAGTGAATCCCTTCTCAGATA 1620  
Db 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAACAAAACAGTGAATCCCTTCTCAGATA 1620  
QY 1621 CTGGGACTAACAGCTTCACTGTGTTGAGGGGCCACGAAGCAAGCAAAAGCTGAAGAA 1680  
Db 1621 CTGGGACTAACAGCTTCACTGTGTTGAGGGGCCACGAAGCAAGCAAAAGCTGAAGAA 1680  
QY 1681 CACATTTTAAACTATGTGCCAATCTTGATGCAATTATTAACCTGTAAAGAAAAAC 1740

Db 1680 CACATTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACTGTTAAAGAAAAC 1739  
Qy 1741 CTTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAAGTTGCCATTGCAGATGCA 1800  
Db 1740 CTTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAAGTTGCCATTGCAGATGCA 1799  
Qy 1801 ACCAGAGAAGATCCATTTTAAAGAGAACTTTAGAAATATGACAGGATTCAACTTAT 1860  
Db 1800 ACCAGAGAAGATCCATTTTAAAGAGAACTTTAGAAATATGACAGGATTCAACTTAT 1859  
Qy 1861 TGTCAATAGATCCAAATGTCAGATTTTGGAACTCAACCCCTATGAACATGGGCCATTCAA 1920  
Db 1860 TGTCAATAGATCCAAATGTCAGATTTTGGAACTCAACCCCTATGAACATGGGCCATTCAA 1919  
Qy 1921 ATGAAAAAAGCTGCAAAAAAGGAAATCGAAAGACGTGTTGTGCGAGAACATTGTG 1980  
Db 1920 ATGAAAAAAGCTGCAAAAAAGGAAATCGAAAGACGTGTTGTGCGAGAACATTGTG 1979  
Qy 1981 AGGAAGTACAATGAGGCCCTACAATTAATGACACAATTGCAATGATGATGCGTATACT 2040  
Db 1980 AGGAAGTACAATGAGGCCCTACAATTAATGACACAATTGCAATGATGATGCGTATACT 2039  
Qy 2041 CATCTTGAACCTTTCTATATATGAAGAGAAAGATAGAAGTTTGCAGTCAATAGAATGAT 2100  
Db 2040 CATCTTGAACCTTTCTATATATGAAGAGAAAGATAGAAGTTTGCAGTCAATAGAATGAT 2099  
Qy 2101 AGTGATGAGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
Db 2100 AGTGATGAGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2159  
Qy 2161 AAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGAACCAAT 2220  
Db 2160 AAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGAACCAAT 2219  
Qy 2221 AAAATGTTGAAAAAGCTGCTGAAAAACCCAGAAATATGAAATGAAAGCTGACCAATTA 2280  
Db 2220 AAAATGTTGAAAAAGCTGCTGAAAAACCCAGAAATATGAAATGAAAGCTGACCAATTA 2279  
Qy 2281 AGAAATACCATATATGAGCAATATATCTAGACTGAGGAATCAGCAGCAGGAATATCTTT 2340  
Db 2280 AGAAATACCATATATGAGCAATATATCTAGACTGAGGAATCAGCAGCAGGAATATCTTT 2339  
Qy 2341 ACAAAAAACAGCAGAGTGCATATGCGCTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400  
Db 2340 ACAAAAAACAGCAGAGTGCATATGCGCTTCCAGTGGATTACTGAAAAATGAAAAATTT 2399  
Qy 2401 GCTGAAGTAGAGTCAAAGCCCAACCATCTGATGGAGCTGACACAGCAGTGAAGTTCAAA 2460  
Db 2400 GCTGAAGTAGAGTCAAAGCCCAACCATCTGATGGAGCTGACACAGCAGTGAAGTTCAAA 2459  
Qy 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAATTTCCGACTGGAATAATCAAT 2520  
Db 2460 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAATTTCCGACTGGAATAATCAAT 2519  
Qy 2521 CTGCTTATCGCTACCAAGTGGCAGAGAAAGCTTGATATTAAGAAATGTAAACATTGTT 2580  
Db 2520 CTGCTTATCGCTACCAAGTGGCAGAGAAAGCTTGATATTAAGAAATGTAAACATTGTT 2579  
Qy 2581 ATCCGTATGCTCTGTCACCAATGAATAGCCATGCTCCAGGCCCGTGTGAGCCAGA 2640  
Db 2580 ATCCGTATGCTCTGTCACCAATGAATAGCCATGCTCCAGGCCCGTGTGAGCCAGA 2639  
Qy 2641 GCTGATGAGAGCACCTAGCTCCTGTTGCTCAGAGTGTTCAGAGATTGCAACATGAG 2700  
Db 2640 GCTGATGAGAGCACCTAGCTCCTGTTGCTCAGAGTGTTCAGAGATTGCAACATGAG 2699  
Qy 2701 ACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTGTCTCAAAATATG 2760  
Db 2700 ACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTGTCTCAAAATATG 2759  
Qy 2761 AAACCAAGAGATGTCTCATTAAGATTTTGAATTAACAGATGCAAAAGTAAATGAAAAAG 2820  
Db 2760 AAACCAAGAGATGTCTCATTAAGATTTTGAATTAACAGATGCAAAAGTAAATGAAAAAG 2819

Db 2760 AAACCAAGAGATGTCTCATTAAGATTTTGAATTAACAGATGCAAAAGTAAATGAAAAAG 2819  
Qy 2821 AAAATGAAACACAGAGAAATATTTGCCAGCATTTACAGAAATTAACCATCACTAATACT 2880  
Db 2820 AAAATGAAACACAGAGAAATATTTGCCAGCATTTACAGAAATTAACCATCACTAATACT 2879  
Qy 2881 TTCCTTGCAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAG 2940  
Db 2880 TTCCTTGCAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAG 2939  
Qy 2941 AAAATGCATCAGCTCAATATGACCCCGAATTTCAAGAACTTTAATGTAAGAGAAAAAC 3000  
Db 2940 AAAATGCATCAGCTCAATATGACCCCGAATTTCAAGAACTTTAATGTAAGAGAAAAAC 2999  
Qy 3001 AAAGCACTGCAAAAGAGTGTGCCGACTATCAATTAATGTAAGAAATCATCTGCAAAATGT 3060  
Db 3000 AAAGCACTGCAAAAGAGTGTGCCGACTATCAATTAATGTAAGAAATCATCTGCAAAATGT 3059  
Qy 3061 GGCCAGGCTTGGGGAACAATGATGTGTGACAAAGGCTTAGATTGCTCTCAAAATA 3120  
Db 3060 GGCCAGGCTTGGGGAACAATGATGTGTGACAAAGGCTTAGATTGCTCTCAAAATA 3119  
Qy 3121 AGGAATTTGTAGTGTGTTTCAAAAAATTAATTCACAAAGAAACATACAAAAAGTGGGTA 3180  
Db 3120 AGGAATTTGTAGTGTGTTTCAAAAAATTAATTCACAAAGAAACATACAAAAAGTGGGTA 3179  
Qy 3181 GAATTACCTATCAATTTCCCAATCTTGACTATTCAGAAATGCTGTTAATTAGTATGAG 3240  
Db 3180 GAATTACCTATCAATTTCCCAATCTTGACTATTCAGAAATGCTGTTAATTAGTATGAG 3239  
Qy 3241 GATTAGCACTTGATTGAAGATTTCTTAAATACTATCAGTTAAACATTTAATATGATTA 3300  
Db 3240 GATTAGCACTTGATTGAAGATTTCTTAAATACTATCAGTTAAACATTTAATATGATTA 3299  
Qy 3301 TGATTAATGTATTCATTTATGCTACAGAACTGACATAGAATCAATAAATGATGTTTAA 3360  
Db 3300 TGATTAATGTATTCATTTATGCTACAGAACTGACATAGAATCAATAAATGATGTTTAA 3359  
Qy 3361 CTCTG 3365  
Db 3360 CTCTG 3364

RESULT 10  
ADQ22237 standard; DNA; 3668 BP.  
ADQ22237;  
26-AUG-2004 (first entry)  
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5057.  
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;  
KW ds.  
XX Homo sapiens.  
PN WO2004048938-A2.  
XX 10-JUN-2004.  
PD 26-NOV-2003; 2003WO-US038193.  
PF 26-NOV-2002; 2002US-0429739P.  
PR (PROT-) PROTEIN DESIGN LABS INC.  
PA Aziz N, Ginsburg WM, Zlotnick A;  
XX WPI; 2004-441208/41.  
DR Early detection of soft tissue sarcoma comprises determining expression  
PT

PT of a gene in a first soft tissue sample and a normal soft tissue sample  
PT and comparing the gene expression, also useful in treating soft tissue  
PT sarcoma.

XX  
PS Example 2; SEQ ID NO 5057; 210bp; English.

XX  
CC The invention relates to a novel method for detecting soft tissue sarcoma  
CC which comprises obtaining a first soft tissue sample from an individual  
CC and a normal soft tissue sample from the same or different individual,  
CC determining the expression of a gene in both samples and comparing the  
CC expression of the gene in both soft tissue samples, where a higher level  
CC of protein expression in the first soft tissue sample indicates the  
CC presence of soft tissue sarcoma. The method of the invention has  
CC cytosstatic applications and may be useful for detecting soft tissue  
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic  
CC acid sequences may be useful in diagnostic and screening applications.  
CC The current sequence is that of a human soft tissue sarcoma-upregulated  
CC DNA of the invention. The current sequence is not shown within the  
CC specification per se but was submitted in CD format by the inventor.

XX  
SQ Sequence 3668 BP; 1215 A; 730 C; 811 G; 872 T; 0 U; 40 Other;

Query Match 99.2%; Score 3339.2; DB 12; Length 3668;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 3341; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 GCGCGCCGGCCTGAGAGCCCTGTGACAACCTCGTCATTGTCCAGGCACAGCGGTAGAC 60  
DB 1 GCGCGCCGGCCTGAGAGCCCTGTGACAACCTCGTCATTGTCCAGGCACAGCGGTAGAC 60  
QY 61 CTTGCTTCTTAAGTGGGACGGACAGCGGACGACATTTCACCTGTCCCGCAGACAA 120  
DB 61 CTTGCTTCTTAAGTGGGACGGGACGACAGCGGACGACATTTCACCTGTCCCGCAGACAA 120  
QY 121 CAGACCATCTGCTTGGGAAACCCCTCTCCCTTCTCTGAGAAGAAAGATGTCGAATGG 180  
DB 121 CAGACCATCTGCTTGGGAAACCCCTCTCCCTTCTCTGAGAAGAAAGATGTCGAATGG 180  
QY 181 TATTTCCACAGACGAGAAATTCGCTATCTCATCTGCTTCAAGGGCCAGGGTGAATG 240  
DB 181 TATTTCCACAGACGAGAAATTCGCTATCTCATCTGCTTCAAGGGCCAGGGTGAATG 240  
QY 241 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTCTGCTGACAGGTGAAGAG 300  
DB 241 TACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTCTGCTGACAGGTGAAGAG 300  
QY 301 CAGATTCAAGAGACAGTCGCCACCTCCGGAACTGACAGGAGTTGAATGCTGCTGAGC 360  
DB 301 CAGATTCAAGAGACAGTCGCCACCTCCGGAACTGACAGGAGTTGAATGCTGCTGAGC 360  
QY 361 ACCTTGAGAGAGGAGTCTGGCACTTGGTGAACCTCGGGAATTCGTGAGGCCCTCCG 420  
DB 361 ACCTTGAGAGAGGAGTCTGGCACTTGGTGAACCTCGGGAATTCGTGAGGCCCTCCG 420  
QY 421 AGAACCAGGACGCCCTGTGGCCGCCCTACATGAACCTGAGCTCACGGACTTGCCCTCT 480  
DB 421 AGAACCAGGACGCCCTGTGGCCGCCCTACATGAACCTGAGCTCACGGACTTGCCCTCT 480  
QY 481 CCATCGTTTGAGAACGCTCATGATGATATCTCAACTGCTGAACCTCTTCAGCCCACT 540  
DB 481 CCATCGTTTGAGAACGCTCATGATGATATCTCAACTGCTGAACCTCTTCAGCCCACT 540  
QY 541 CTGGTGACAAGCTTCTAGTTAGAGCGTCTTGATTAAGTCATGAGGAGAACTGTTG 600  
DB 541 CTGGTGACAAGCTTCTAGTTAGAGCGTCTTGATTAAGTCATGAGGAGAACTGTTG 600  
QY 601 ACAATTGAAGACAGAAACCGGATGCTGCTGCAGAAAACAATGGAATGAATCAGGTGA 660  
DB 601 ACAATTGAAGACAGAAACCGGATGCTGCTGCAGAAAACAATGGAATGAATCAGGTGA 660  
QY 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAACTGGTCTCTGCATTCTGAATGTT 720  
DB 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAGAAACTGGTCTCTGCATTCTGAATGTT 720

QY 721 CTTGCTCAACACAGAAACAATGAACCTGTGCCAAGAGTTAACAGGCTGTGCTCAGAA 780  
DB 721 NNNNNNNNNNNNNNNNNNNATGAACCTGTGCCAAGAGTTAACAGGCTGTGCTCAGAA 780  
QY 781 AGCAATGCAGAGATTGAGAATTTATCACAGTTGATGTCCTCAAGTGAAGCAACTT 840  
DB 781 AGCAATGCAGAGATTGAGAATTTATCACAGTTGATGTCCTCAAGTGAAGCAACTT 840  
QY 841 CTTTCAACCAAGTTCAAGCCAAATCTGGAAGAGAGTCTGGGCAATGGAATACTCA 900  
DB 841 CTTTCAACCAAGTTCAAGCCAAATCTGGAAGAGAGTCTGGGCAATGGAATACTCA 900  
QY 901 TCAGATCATCTTTGACAGATTCTTGTAGTTTCAAGATCAGACACAAGTTGGCAGAA 960  
DB 901 TCAGATCATCTTTGACAGATTCTTGTAGTTTCAAGATCAGACACAAGTTGGCAGAA 960  
QY 961 GGAAGTGCAGCTGCTTAGATGAAGTCTTGGACATACAGCAATGGGCAGTATTC 1020  
DB 961 GGAAGTGCAGCTGCTTAGATGAAGTCTTGGACATACAGCAATGGGCAGTATTC 1020  
QY 1021 GGCACCATGGGAGAGTATTCAGATGAAGAGATGTGGCAGCAAGCATCCCGAGCCA 1080  
DB 1021 GGCACCATGGGAGAGTATTCAGATGAAGAGATGTGGCAGCAAGCATCCCGAGCCA 1080  
QY 1081 GAATCCAGCTCAGGCTTACCAATGGAAGTGGCCAGCCAGCTTGAAGGGAAGAT 1140  
DB 1081 GAATCCAGCTCAGGCTTACCAATGGAAGTGGCCAGCCAGCTTGAAGGGAAGAT 1140  
QY 1141 ATCATCATCTGCTCCCTACAGGAGTGAAGAAACAGAGTGGCTTTACATTGCCAG 1200  
DB 1141 ATCATCATCTGCTCCCTACAGGAGTGAAGAAACAGAGTGGCTTTACATTGCCAG 1200  
QY 1201 GATCACTTAGACAAAGAAAGAAAGCATCTGAGCCCTGGAAGAGTTAAGTTCTGTCAAT 1260  
DB 1201 GATCACTTAGACAAAGAAAGAAAGCATCTGAGCCCTGGAAGAGTTAAGTTCTGTCAAT 1260  
QY 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGATTCCAACTTTTGAAGAAATGG 1320  
DB 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGATTCCAACTTTTGAAGAAATGG 1320  
QY 1321 TATCGTTTATTTGATTAAGTGTGATACCAACTGAAGAAATATCATTTCCAGAAATGTC 1380  
DB 1321 TATCGTTTATTTGATTAAGTGTGATACCAACTGAAGAAATATCATTTCCAGAAATGTC 1380  
QY 1381 AAGTCTGTATATTAATATACAGTACAGCTCAATCCTTGAAGAACTCCCTTAAACTTG 1440  
DB 1381 AAGTCTGTATATTAATATACAGTACAGCTCAATCCTTGAAGAACTCCCTTAAACTTG 1440  
QY 1441 GAAAATGAGAGATGCTGTGTTCAATTTGACAGCTTTCCCTCATTTATCATGTGATGA 1500  
DB 1441 GAAAATGAGAGATGCTGTGTTCAATTTGACAGCTTTCCCTCATTTATCATGTGATGA 1500  
QY 1501 TGTCACTCACACCAAGAAAGAGCAGTATTAATAATCATGAGGCAATTTGATGACAG 1560  
DB 1501 TGTCACTCACACCAAGAAAGAGCAGTATTAATAATCATGAGGCAATTTGATGACAG 1560  
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACAACCAAGTATTCCTCTCAGATA 1620  
DB 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACAACCAAGTATTCCTCTCAGATA 1620  
QY 1621 CTGGAGCTAACAGCTTCACTGTGTGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
DB 1621 CTGGAGCTAACAGCTTCACTGTGTGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680  
QY 1681 CACATTTTAAACTATGTGCCAATCTTGATGCAATTTACTATTAACAATGTTAAGAAAC 1740  
DB 1681 CACATTTTAAACTATGTGCCAATCTTGATGCAATTTACTATTAACAATGTTAAGAAAC 1740  
QY 1741 CTTGATCAACTGAAGAAACCAATATACAGAGCCATGCAAGAGTTGGCAATGACATGCA 1800  
DB 1741 CTTGATCAACTGAAGAAACCAATATACAGAGCCATGCAAGAGTTGGCAATGACATGCA 1800

QY	1801	ACCAGAGAGATCCATTTAAAGAAA	CTTCTAGAAATAATGACAAGGATTCAACTTAT	1860			
Db	1801	ACCAGAGAGATCCATTTAAAGAAA	CTTCTAGAAATAATGACAAGGATTCAACTTAT	1860			
QY	1861	TGTCAAATGAGTCCAAATGTCA	GATTTTGGAACTCAACCTATGAA	CAATGGGCCATTCAA	1920		
Db	1861	TGTCAAATGAGTCCAAATGTCA	GATTTTGGAACTCAACCTATGAA	CAATGGGCCATTCAA	1920		
QY	1921	ATGAAAAAAAAGCTGC	AAAAAAGGAATCGCA	AAAGACGTGTTGTG	CAGAACATTTG	1980	
Db	1921	ATGAAAAAAAAGCTGC	AAAAAAGGAATCGCA	AAAGACGTGTTGTG	CAGAACATTTG	1980	
QY	1981	AGGAAGTACAATGAGGCCCTA	CAAAATTAATGACACAATTG	CAATGATAGATGCGTAT	ACT	2040	
Db	1981	AGGAAGTACAATGAGGCCCTA	CAAAATTAATGACACAATTG	CAATGATAGATGCGTAT	ACT	2040	
QY	2041	CATCTTGAACCTTCTAT	TAATGAAGAAAGATAAGAGTTG	CAGTCATAGAAAGAT		2100	
Db	2041	CATCTTGAACCTTCTAT	TAATGAAGAAAGATAAGAGTTG	CAGTCATAGAAAGAT		2100	
QY	2101	AGTGATGAGGGTGGTGAT	GAAGTATTTGTGATGGTGAT	GAAGATGAGGATGATTT	TAAG	2160	
Db	2101	AGTGATGAGGGTGGTGAT	GAAGTATTTGTGATGGTGAT	GAAGATGAGGATGATTT	TAAG	2160	
QY	2161	AAACCTTTGAAACTG	ATGAAACAGATGATTTCT	CAATGACTTTATTTT	TGAAAACAAT	2220	
Db	2161	AAACCTTTGAAACTG	ATGAAACAGATGATTTCT	CAATGACTTTATTTT	TGAAAACAAT	2220	
QY	2221	AAAAATGTTGAAAAAG	CTGCTGAAAAAACCAGAA	TATGAAAAATGAAAA	AGCTGACCAATTA	2280	
Db	2221	AAAAATGTTGAAAAAG	CTGCTGAAAAAACCAGAA	TATGAAAAATGAAAA	AGCTGACCAATTA	2280	
QY	2281	AGAAATACCATTAATG	AGCAATATACTAGAGCTG	AGGAATCAGCAGAGGA	ATACTTT	2340	
Db	2281	AGAAATACCATTAATG	AGCAATATACTAGAGCTG	AGGAATCAGCAGAGGA	ATACTTT	2340	
QY	2341	ACAAAAACAGCAGACAG	ATGATATGCGCTTTCC	AGTGAATTA	CTGAAAAATGAAAA	TTT	2400
Db	2341	ACAAAAACAGCAGACAG	ATGATATGCGCTTTCC	AGTGAATTA	CTGAAAAATGAAAA	TTT	2400
QY	2401	GCTGAAGTAGAGTCA	AAAGCCCA	CCATCTGATTGGAGCTG	GACACAGCAGTGA	TTCAAA	2460
Db	2401	GCTGAAGTAGAGTCA	AAAGCCCA	CCATCTGATTGGAGCTG	GACACAGCAGTGA	TTCAAA	2460
QY	2461	CCCATGACACAGAA	TGAACAAAAAGAGTCA	TTAGTAAATTTG	CGCACTG	AAAAATTAAT	2520
Db	2461	CCCATGACACAGAA	TGAACAAAAAGAGTCA	TTAGTAAATTTG	CGCACTG	AAAAATTAAT	2520
QY	2521	CTGCTTATCGCTAC	ACACAGTGGCAGAA	AGGCTCTGATATTA	AAAGATGTAA	CACTTGT	2580
Db	2521	CTGCTTATCGCTAC	ACACAGTGGCAGAA	AGGCTCTGATATTA	AAAGATGTAA	CACTTGT	2580
QY	2581	ATCCGTTATG	GTCTGTCACCAATGA	ATAGCCATGGTCC	AGGCCCGTGTG	CAGCCAGA	2640
Db	2581	ATCCGTTATG	GTCTGTCACCAATGA	ATAGCCATGGTCC	AGGCCCGTGTG	CAGCCAGA	2640
QY	2641	GCTGATGAGAC	CCCTACGTCCTG	TTGCTCACAGTGGTT	CAAGAGTTATCG	AAACATGAG	2700
Db	2641	GCTGATGAGAC	CCCTACGTCCTG	TTGCTCACAGTGGTT	CAAGAGTTATCG	AAACATGAG	2700
QY	2701	ACAGTTAATGATTT	CCGAGAGAAAGATG	ATGATTAAGCTAT	ACATTTGTGTT	CAAAATATG	2760
Db	2701	ACAGTTAATGATTT	CCGAGAGAAAGATG	ATGATTAAGCTAT	ACATTTGTGTT	CAAAATATG	2760
QY	2761	AAACGAGAGAGTATG	CTCATTAAGATTTT	GGAATTA	CAAGTGC	AAAGTATATG	2820
Db	2761	AAACGAGAGAGTATG	CTCATTAAGATTTT	GGAATTA	CAAGTGC	AAAGTATATG	2820
QY	2821	AAAAATGA	AAAAACCAAGAAATAT	TGCCAAGCATTA	CAAGAAATTA	CCCATCACTA	2880
Db	2821	AAAAATGA	AAAAACCAAGAAATAT	TGCCAAGCATTA	CAAGAAATTA	CCCATCACTA	2880
QY	2881	TTCC	TTTGCAAAAACTG	CAGTGTGCTAG	CCCTGTTCTG	GGGAAAGATAT	2940
Db	2881	TTCC	TTTGCAAAAACTG	CAGTGTGCTAG	CCCTGTTCTG	GGGAAAGATAT	2940

Db 2881 TTCCTTGC AAAA AACTGCAGTGTGCTAGCCTGTCTGGGAAGATATCCATGTAATTGAG 2940

QY 2941 AAAATGCATCAGCTCAATATATGACCCCGAAGATTCAAGGACTTTACATTGTAAAGAAAC 3000

Db 2941 AAAATGCATCAGCTCAATATATGACCCCGAAGATTCAAGGACTTTACATTGTAAAGAAAC 3000

QY 3001 AAAGCACTGC AAAAGAAAGTGTGCCGACTATCAAAATAAATGCTGAATCATCTGC AAATGT 3060

Db 3001 AAAGCACTGC AAAAGAAAGTGTGCCGACTATCAAAATAAATGCTGAATCATCTGC AAATGT 3060

QY 3061 GGCCAGGCTTGGGGNACAATGATGTGCACAAAGGCTTAGATTTGCCCTGTCTCAAAATA 3120

Db 3061 GGCCAGGCTTGGGGNACAATGATGTGCACAAAGGCTTAGATTTGCCCTGTCTCAAAATA 3120

QY 3121 AGGAATTTTGTAGTGTCTTTC AAAAATTAATTCACAAGAACAATACAAAAGTGGGTA 3180

Db 3121 AGGAATTTTGTAGTGTCTTTC AAAAATTAATTCACAAGAACAATACAAAAGTGGGTA 3180

QY 3181 GAATTACCTATCACATTTCCCAATCTTGACTAATTCAGAATGCTGTTATTAGTGATGAG 3240

Db 3181 GAATTACCTATCACATTTCCCAATCTTGACTAATTCAGAATGCTGTTATTAGTGATGAG 3240

QY 3241 GATTAGCACTTGATTGAAGATTCTTTAAAAATACTATCAGTTAAACATTTAATATGATTA 3300

Db 3241 GATTAGCACTTGATTGAAGATTCTTTAAAAATACTATCAGTTAAACATTTAATATGATTA 3300

QY 3301 TGATTAATGTATTCATTATGTCTACAGAACTGCATAAAGATCAATAAAATGATTGTTTAA 3360

Db 3301 TGATTAATGTATTCATTATGTCTACAGAACTGCATAAAGATCAATAAAATGATTGTTTAA 3360

QY 3361 CTCTG 3365

Db 3361 CTCTG 3365

RESULT	11
ABA04908	
ID	ABA04908 standard; cDNA; 3372 BP.
XX	
AC	ABA04908;
XX	
DT	05-MAR-2002 (first entry)
XX	
DE	Human RNA helicase RH16 coding sequence.
XX	
KW	Human; RH16; RNA helicase; cytosratic; virucide; anti-HIV;
KW	Immunosuppressive; immunostimulatory; antirheumatic; antiarthritic;
KW	antiartherosclerotic; osteopathic; antidiabetic; hepatrotropic;
KW	antiinflammatory; cancer; infection; HIV; hepatitis; genetic disease;
KW	autoimmune disease; graft rejection; vaccine; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	155..3232
FT	/*tag= a
FT	/product= "Human RH16"
XX	
PN	WO200185955-A1.
XX	
PD	15-NOV-2001.
XX	
Pf	11-MAY-2001; 2001WO-FR001441.
XX	
PR	11-MAY-2000; 2000FR-00006030.
XX	
PA	(ISTA-) ISTAC.
PA	(INSP ) INST PASTEUR LILLE.
XX	
PI	Bahr G, Cocude C, Capron A;
XX	
DR	WPI; 2002-082898/11.

DR P-PSDB; AAM47798.

XX New polypeptide, useful for treating and diagnosing cancer or  
PT inflammation, and drug screening, comprises a human polynucleotide  
PT homologous to RNA helicase.

XX Claim 7; Page 85-89; 114pp; French.

CC The present sequence is the coding sequence for human RH16. RH16 is a  
CC 116kDa protein and has homology to RNA helicases (DEXH box). RH16 and  
CC its coding sequence are useful for treating cancer; acute or chronic  
CC infections (especially by HIV or hepatitis B or C); inherited genetic  
CC diseases; (auto)immune diseases (particularly rheumatism, arthritis,  
CC arteriosclerosis, osteoporosis and diabetes, but many others listed) and  
CC to prevent graft rejection. RH16 and its coding sequence are also useful  
CC for inducing, or increasing, the immune response to a vaccine

XX Sequence 3372 BP; 1157 A; 637 C; 748 G; 830 T; 0 U; 0 Other;

Query Match 99.2%; Score 3338.8; DB 6; Length 3372;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3343; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 15 GAGCCCTGTGACAACTCGTCATTGTTCAGGACAGCGGTAGACCCTCTCTCTAAG 74  
Db 1 GGGCCCTGTGACAACTCGTCATTGTTCAGGACAGCGGTAGACCCTCTCTCTAAG 60  
QY 75 TGGGCACGGGACAGCGGACGACATTTACCTGTCCCGACAGACAACACCATCTGCT 134  
Db 61 TGGGCACGGGACAGCGGACGACATTTACCTGTCCCGACAGACAACACCATCTGCT 120  
QY 135 TGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTGGAATGGGTATTCACAGACGA 194  
Db 121 TGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTGGAATGGGTATTCACAGACGA 180  
QY 195 GAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAATAATGTACATCCAGGTGA 254  
Db 181 GAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAATAATGTACATCCAGGTGA 240  
QY 255 GCCTGTGCTGACTACCTGACTCTTCTGCTGCTGAGAGGTGAAGAGACGATTTCAGAGAC 314  
Db 241 GCCTGTGCTGACTACCTGACTCTTCTGCTGCTGAGAGGTGAAGAGACGATTTCAGAGAC 300  
QY 315 AGTCGCCACCTCCGGGAACATGACGACGAGTTGAACCTGCTGAGCACCCTTGAGAAGGG 374  
Db 301 AGTCGCCACCTCCGGGAACATGACGACGAGTTGAACCTGCTGAGCACCCTTGAGAAGGG 360  
QY 375 AGTCTGGCACCTTGGTGGACTCGGGAATTCTGTGAGAGCCCTCCGGAGAACCGGCAGCCC 434  
Db 361 AGTCTGGCACCTTGGTGGACTCGGGAATTCTGTGAGAGCCCTCCGGAGAACCGGCAGCCC 420  
QY 435 TCTGGCCGCGCTACATGAACCTGAGCTACGGACTTGCCCTCTCCATCGTTTGAGAA 494  
Db 421 TCTGGCCGCGCTACATGAACCTGAGCTACGGACTTGCCCTCTCCATCGTTTGAGAA 480  
QY 495 CGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTTCAGCCCACTCTGTGGAAGAGCT 554  
Db 481 CGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTTCAGCCCACTCTGTGGAAGAGCT 540  
QY 555 TCTAGTTAGAGACGCTTGGATTAAGTCATGAGAGAGAACTGTTGACAATTGAAGACAG 614  
Db 541 TCTAGTTAGAGACGCTTGGATTAAGTCATGAGAGAGAACTGTTGACAATTGAAGACAG 600  
QY 615 AAACCGGATTGCTGCTGACAGAAAACAATGGAATGAATCAAGGTGAAGAGACTTAATA 674  
Db 601 AAACCGGATTGCTGCTGACAGAAAACAATGGAATGAATCAAGGTGAAGAGACTTAATA 660  
QY 675 AAGGATGTGCAGAAAGAAAAGTGTCTCTGCAATTTCTGAATGTTCTTCTCAACAGAG 734  
Db 661 AAGGATGTGCAGAAAGAAAAGTGTCTCTGCAATTTCTGAATGTTCTTCTCAACAGAG 720  
QY 735 AAACAATGAACCTGTCCAGAGTTAACAGGCTGTGATGCTCAGAAAGCAATGCAGAGAT 794  
Db 721 AAACAATGAACCTGTCCAGAGTTAACAGGCTGTGATGCTCAGAAAGCAATGCAGAGAT 780

QY 795 TGAGATTTATCACAAGTTGATGTCCTCAAGTGAAGAGCAACTCTTCAACCACAGT 854  
Db 781 TGAGATTTATCACAAGTTGATGTCCTCAAGTGAAGAGCAACTCTTCAACCACAGT 840  
QY 855 TCAGCCAAATCTGGAGAAAGGAGTCTGGGCGCATGGAGATACTCATCAGATCATCTTT 914  
Db 841 TCAGCCAAATCTGGAGAAAGGAGTCTGGGCGCATGGAGATACTCATCAGATCATCTTT 900  
QY 915 TGCAGATTTCTGTAGTTTTCAGAAATCAGACACAAGTTGGCAGAAAGAGTGCAGCTG 974  
Db 901 TGCAGATTTCTGTAGTTTTCAGAAATCAGACACAAGTTGGCAGAAAGAGTGCAGCTG 960  
QY 975 CTTAGATGAAGTCTTGACATTAACAGCAATGGGCGAGTGAATTCAGGCAACCATGGGAAG 1034  
Db 961 CTTAGATGAAGTCTTGACATTAACAGCAATGGGCGAGTGAATTCAGGCAACCATGGGAAG 1020  
QY 1035 TGATTCAGATGAAGAGATGTGGCAGCAGACATCCCCGAGCCAGAACTCCAGCTCAG 1094  
Db 1021 TGATTCAGATGAAGAGATGTGGCAGCAGACATCCCCGAGCCAGAACTCCAGCTCAG 1080  
QY 1095 GCCTTACCAATGGAAGTTGCCACGACCTTGGAAGGGAAGATATCATCATCTGCCT 1154  
Db 1081 GCCTTACCAATGGAAGTTGCCACGACCTTGGAAGGGAAGATATCATCATCTGCCT 1140  
QY 1155 CCCTACAGGAGTGGAAAAACCAAGATGGCTTTACATTTGCCAAGATCACTTAGACAA 1214  
Db 1141 CCCTACAGGAGTGGAAAAACCAAGATGGCTTTACATTTGCCAAGATCACTTAGACAA 1200  
QY 1215 GAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGTACTGCTAGT 1274  
Db 1201 GAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGTACTGCTAGT 1260  
QY 1275 TGAACAGCTCTCCGCAAGAGATTCCAACTTTTGAAGAAATGGTATCGTTATTGG 1334  
Db 1261 TGAACAGCTCTCCGCAAGAGATTCCAACTTTTGAAGAAATGGTATCGTTATTGG 1320  
QY 1335 ATTAAGTGTGATACCACTGAAGTGAATAATTCACAGAAATTGTCAAGTCCGTGATAT 1394  
Db 1321 ATTAAGTGTGATACCACTGAAGTGAATAATTCACAGAAATTGTCAAGTCCGTGATAT 1380  
QY 1395 TATTATCAGTACAGTCAAAATCCTTGAATACTCCCTTTAACTTGAATAATGAGAAGA 1454  
Db 1381 TATTATCAGTACAGTCAAAATCCTTGAATACTCCCTTTAACTTGAATAATGAGAAGA 1440  
QY 1455 TGCTGTGTTCAATTGTCAAGACTTTTCCCTCATTTATCATTTGATGAATGTCATCACACCA 1514  
Db 1441 TGCTGTGTTCAATTGTCAAGACTTTTCCCTCATTTATCATTTGATGAATGTCATCACACCA 1500  
QY 1515 CAAAGAACAGTGTATAATAACATCATGAGCAATTATTGATGAGAAAGTGAACAA 1574  
Db 1501 CAAAGAACAGTGTATAATAACATCATGAGCAATTATTGATGAGAAAGTGAACAA 1560  
QY 1575 TAGACTCAAGAAAGAAAACAACCAAGTATCCCTTCTCAGATACTGGAGCTAACAGC 1634  
Db 1561 TAGACTCAAGAAAGAAAACAACCAAGTATCCCTTCTCAGATACTGGAGCTAACAGC 1620  
QY 1635 TTCACTGTGTGAGAGGGGCCACGAGAACCCAAAGCTGAAGAACACATTTTAAACT 1694  
Db 1621 TTCACTGTGTGAGAGGGGCCACGAGAACCCAAAGCTGAAGAACACATTTTAAACT 1680  
QY 1695 ATGTGCCAATCTTGATGATTTACTATTTAAACTGTTTAAGAAAACTTGATCAACTGAA 1754  
Db 1681 ATGTGCCAATCTTGATGATTTACTATTTAAACTGTTTAAGAAAACTTGATCAACTGAA 1740  
QY 1755 AAACCAATACAGAGCCATGCAAGAGTTGCCATTCAGATGCAACAGAGAGATCC 1814  
Db 1741 AAACCAATACAGAGCCATGCAAGAGTTGCCATTCAGATGCAACAGAGAGATCC 1800  
QY 1815 ATTTAAGAGAACTTCTAGAATAATGACAGAGATTCAAACCTTATTTGTCAAAATGAGTCC 1874  
Db 1801 ATTTAAGAGAACTTCTAGAATAATGACAGAGATTCAAACCTTATTTGTCAAAATGAGTCC 1860

```
OY 1875 AATGTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAATGGAAAAAAGC 1934
Db 1861 AATGTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAATGGAAAAAAGC 1920
OY 1935 TGCAAAAAAGAAATCGCAAGAAGCTGTTGTGCAAAACATTGGAGAGTACAATGA 1994
Db 1921 TGCAAAAAAGAAATCGCAAGAAGCTGTTGTGCAAAACATTGGAGAGTACAATGA 1980
OY 1995 GGGCCCTACAATTAATGACACAAATTCGAATGATAGATGCGTATACTCATCTTGAACCTT 2054
Db 1981 GGGCCCTACAATTAATGACACAAATTCGAATGATAGATGCGTATACTCATCTTGAACCTT 2040
OY 2055 CTATTAATGAAGAGAAAGATAAGATTGGCAGTCATAGAAGATGATAGTAGAGGGTGG 2114
Db 2041 CTATTAATGAAGAGAAAGATAAGATTGGCAGTCATAGAAGATGATAGTAGAGGGTGG 2100
OY 2115 TGATGATGAGTATTGTGATGCTGATGAGATGAGATGATTAAAGAAACCTTGAACCT 2174
Db 2101 TGATGATGAGTATTGTGATGCTGATGAGATGAGATGATTAAAGAAACCTTGAACCT 2160
OY 2175 GGATGAAAACAGATAGATTTCTCATGACTTATTTTGAACCAATAAATGTTGAAAG 2234
Db 2161 GGATGAAAACAGATAGATTTCTCATGACTTATTTTGAACCAATAAATGTTGAAAG 2220
OY 2235 GCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAATTAAGAAATACCATAAT 2294
Db 2221 GCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAATTAAGAAATACCATAAT 2280
OY 2295 GGAGCAATATACTAGAGCTGAGAAATCAGACAGAGAAATACTTTACAAAAACAGACA 2354
Db 2281 GGAGCAATATACTAGAGCTGAGAAATCAGACAGAGAAATACTTTACAAAAACAGACA 2340
OY 2355 GAGTGCATATGCGCTTCCAGTGTGATTACTGAAAAATGAAAAATTTGCTGAAGTAGAGT 2414
Db 2341 GAGTGCATATGCGCTTCCAGTGTGATTACTGAAAAATGAAAAATTTGCTGAAGTAGAGT 2400
OY 2415 CAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAATTCAAAACCATGACACAGAA 2474
Db 2401 CAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAATTCAAAACCATGACACAGAA 2460
OY 2475 TGAACAAAAAAGAGTCATTAGTAATTTCCGACCTGGAATAATCAATCTGCTTAATCGCTAC 2534
Db 2461 TGAACAAAAAAGAGTCATTAGTAATTTCCGACCTGGAATAATCAATCTGCTTAATCGCTAC 2520
OY 2535 CACAGTGGCAGAGAAGGCTCGATATTAAAGAATGTAACTTGTATCCGTTATGCTCT 2594
Db 2521 CACAGTGGCAGAGAAGGCTCGATATTAAAGAATGTAACTTGTATCCGTTATGCTCT 2580
OY 2595 CGTCAACCAATGAATAGCCATGCTCCAGGCCCGTGTGAGCCAGAGCTGATGAGAGCAC 2654
Db 2581 CGTCAACCAATGAATAGCCATGCTCCAGGCCCGTGTGAGCCAGAGCTGATGAGAGCAC 2640
OY 2655 CTACGTCCTGCTGCTCTCAAGTGTTCAGGAGTTATGAAACATGAGACAGTTAATGATT 2714
Db 2641 CTACGTCCTGCTGCTCTCAAGTGTTCAGGAGTTATGAAACATGAGACAGTTAATGATT 2700
OY 2715 CCGAGAGAAGATGATGTATTAAGCTATACATTGTGTTCAAAATATGAAACCAAGAGAGTA 2774
Db 2701 CCGAGAGAAGATGATGTATTAAGCTATACATTGTGTTCAAAATATGAAACCAAGAGAGTA 2760
OY 2775 TGTCTCAATGAATTTTGAATTAAGATGCAAAAGTATATAAGAAAAAGAAAAATGAAACCAA 2834
Db 2761 TGTCTCAATGAATTTTGAATTAAGATGCAAAAGTATATAAGAAAAAGAAAAATGAAACCAA 2820
OY 2835 GAGAAATATTTGCCAAGCATTAAGAATAAACCATCACTAATACTTTCTTGCAAAAA 2894
Db 2821 GAGAAATATTTGCCAAGCATTAAGAATAAACCATCACTAATACTTTCTTGCAAAAA 2880
OY 2895 CTGCAAGTGTCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAGAAAAATGATCAGCT 2954
Db 2881 CTGCAAGTGTCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAGAAAAATGATCAGCT 2940
```

```
OY 2955 CAATATGACCCCGAATTCAGGAACCTTTACATTTGTAAGAGAAAAAGCACTGCAAAA 3014
Db 2941 CAATATGACCCCGAATTCAGGAACCTTTACATTTGTAAGAGAAAAAGCACTGCAAAA 3000
OY 3015 GAAAGTGTCCGACTATCAATTAATGTTGTAATCATCTGCAAAATGTGGCCAGGCTGGGG 3074
Db 3001 GAAAGTGTCCGACTATCAATTAATGTTGTAATCATCTGCAAAATGTGGCCAGGCTGGGG 3060
OY 3075 AACATGATGTGTGACAAAGGCTTAGATTGCTGTCTCAAAATAGAAATTTGTAGT 3134
Db 3061 AACATGATGTGTGACAAAGGCTTAGATTGCTGTCTCAAAATAGAAATTTGTAGT 3120
OY 3135 GGTTCCTCAAAATTAATCAACAAAGAAACATAAAGTGGTGAATTAACCTATCAC 3194
Db 3121 GGTTCCTCAAAATTAATCAACAAAGAAACATAAAGTGGTGAATTAACCTATCAC 3180
OY 3195 ATTTCCCAATCTTGACTATTCAGATGCTGTTTATTTAGTATGAGATTAAGCACTTGAT 3254
Db 3181 ATTTCCCAATCTTGACTATTCAGATGCTGTTTATTTAGTATGAGATTAAGCACTTGAT 3240
OY 3255 TGAAGATTTCTTTAAATTAATCTATCAGTTAAACATTTAATATGATTAATGATTC 3314
Db 3241 TGAAGATTTCTTTAAATTAATCTATCAGTTAAACATTTAATATGATTAATGATTC 3300
OY 3315 ATTATGCTACAGAACTGACATTAAGAATCAATTAATGATTTGTTACTCT 3364
Db 3301 ATTATGCTACAGAACTGACATTAAGAATCAATTAATGATTTGTTACTCT 3350
```

```
RESULT 12
AD30823
ID AD30823 standard; cDNA; 3446 BP.
XX
AC AD30823;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel cDNA sequence, SEQ ID NO:905.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
XX
DR P-PSDB; AD31794.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 1; SEQ ID NO 905; 1185bp; English.
```

XX The invention relates to 971 novel human cDNA sequences (ADC29919-  
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
CC invention also relates to nucleic acid sequences over 99% identical with  
CC the novel human cDNAs. The invention additionally encompasses expression  
CC vectors and host cells comprising a nucleic acid of the invention; the  
CC recombinant production of a polypeptide of the invention; an antibody  
CC against a polypeptide of the invention; a method of detecting  
CC polynucleotides or polypeptides of the invention; and methods of  
CC identifying a compound which binds to a polypeptide of the invention. The  
CC invention further discloses methods of preventing, treating or  
CC ameliorating a medical condition; kits comprising polynucleotide probes  
CC and/or monoclonal antibodies for carrying out the methods of the  
CC invention; methods for the identification of compounds that modulate the  
CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention  
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33394). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are  
CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anaemia, platelet  
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a specifically  
CC claimed human cDNA sequence of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 3446 BP; 1184 A; 649 C; 772 G; 841 T; 0 U; 0 Other;

Query Match 98.5%; Score 3316.2; DB 10; Length 3446;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3321; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 36 CATGTGACGGCACAGAGCGGTAGACCCCTGTTCTTAAGTGGGACGGGACAGCGGCACG 95  
DB 104 CACTTTCAGGCACAGAGCGGTAGACCCCTGTTCTTAAGTGGGACGGGACAGCGGCACG 163  
QY 96 CACATTTTCACCTGTCCCGCAGACAACAGCACCATCTGCTGGAGAACCTCTCCCTTCT 155  
DB 164 CACATTTTCACCTGTCCCGCAGACAACAGCACCATCTGCTGGAGAACCTCTCCCTTCT 223  
QY 156 CTGAGAAAGAAAGATGTGAATGGGTATTCCACAGACGAGAATTTCCGTAATCTCATCTC 215  
DB 224 CTGAGAAAGAAAGATGTGAATGGGTATTCCACAGACGAGAATTTCCGTAATCTCATCTC 283  
QY 216 GTGCTTCAGGGCCAGGGGTAAATATCATCCAGTGGAGCCTGTGCTGAATACTGAC 275  
DB 284 GTGCTTCAGGGCCAGGGGTAAATATCATCCAGTGGAGCCTGTGCTGAATACTGAC 343  
QY 276 CTTTCTGCTGACAGAGGTGAAGAGCAGATTCAAGAGCAAGTCCCACTCCGGGAACAT 335  
DB 344 CTTTCTGCTGACAGAGGTGAAGAGCAGATTCAAGAGCAAGTCCCACTCCGGGAACAT 403  
QY 336 GCAGGCAATTGAATGCTGCTGAGACCTTTGAGAAAGGAGTCTGGACACCTTGTTGAGC 395  
DB 404 GCAGGCAATTGAATGCTGCTGAGACCTTTGAGAAAGGAGTCTGGACACCTTGTTGAGC 463  
QY 396 TCGGGAATTCGTGAGGCGCTCCGAGAACCCGACAGCCCTCTGGCCCGCCCTACATGAA 455  
DB 464 TCGGGAATTCGTGAGGCGCTCCGAGAACCCGACAGCCCTCTGGCCCGCCCTACATGAA 523  
QY 456 CCCTGAGCTACGAGACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCA 515  
DB 524 CCCTGAGCTACGAGACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCA 583  
QY 516 ACTGCTGAACCTCCTTCAAGCCCACTCTGTGAGCAAGCTTCTAGTTAGAGACGTCTTGA 575

DB 584 ACTGCTGAACCTCCTTACGCCCACTCTGTGAGCAAGCTTCTAGTTAGAGACGTCTTGA 643  
QY 576 TAAAGTCATGAGAGGAGAACTGTTGACAAATGGAAGACAGAAACCAGATTGCTGTCAGA 635  
DB 644 TAAAGTCATGAGAGGAGAACTGTTGACAAATGGAAGACAGAAACCAGATTGCTGTCAGA 703  
QY 636 AAACAAATGMAATGAATCAGGTGTAAGAGACTATAAAAAGATTGTGCAGAAAGAAA 695  
DB 704 AAACAAATGMAATGAATCAGGTGTAAGAGACTATAAAAAGATTGTGCAGAAAGAAA 763  
QY 696 CTGTTCTCTGCAATTTCTGAATGTTCTGTCAAACAGAAACAAATGAACTGTCCAAGA 755  
DB 764 CTGTTCTCTGCAATTTCTGAATGTTCTGTCAAACAGAAACAAATGAACTGTCCAAGA 823  
QY 756 GTTAAACAGGCTCTGATTTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGA 815  
DB 824 GTTAAACAGGCTCTGATTTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGA 883  
QY 816 TGGTCTCAAGTGAAGAGCAACTCTTTCAACCAACAGTTCAGCCAAATCTGGAGAAAGA 875  
DB 884 TGGTCTCAAGTGAAGAGCAACTCTTTCAACCAACAGTTCAGCCAAATCTGGAGAAAGA 943  
QY 876 GGTCTGGGCAATGAGAAATTAATCAATCAGATCATCTTTTGACAGATTCTTGTAGTTTC 935  
DB 944 GGTCTGGGCAATGAGAAATTAATCAATCAGATCATCTTTTGACAGATTCTTGTAGTTTC 1003  
QY 936 AGAATCAGACACAAGTTTGGCAGAAAGAGTGTACGCTGTTAGTGAAGTCTTGACA 995  
DB 1004 AGAATCAGACACAAGTTTGGCAGAAAGAGTGTACGCTGTTAGTGAAGTCTTGACA 1063  
QY 996 TAAACAGCAACATGGGCGATGATTCAAGGACCATGGGAAGTGAATGATGAAGAAATGT 1055  
DB 1064 TAAACAGCAACATGGGCGATGATTCAAGGACCATGGGAAGTGAATGATGAAGAAATGT 1123  
QY 1056 GGCAAGCAAGACATCCCCGGAGCCAGAACTCCAGCTCAGGCTTAACCAATGGAAGTTGC 1115  
DB 1124 GGCAAGCAAGACATCCCCGGAGCCAGAACTCCAGCTCAGGCTTAACCAATGGAAGTTGC 1183  
QY 1116 CCAGCCAGCCTTGGAAAGGAAATATATCATCTGCTCCCTACAGGAGAGTGAAGAAAAC 1175  
DB 1184 CCAGCCAGCCTTGGAAAGGAAATATATCATCTGCTCCCTACAGGAGAGTGAAGAAAAC 1243  
QY 1176 CAGAGTGGCTGTTACATTGCAAGGATCACTTAGACAAGAAAGAAAAGCATCTGAGCC 1235  
DB 1244 CAGAGTGGCTGTTACATTGCAAGGATCACTTAGACAAGAAAGAAAAGCATCTGAGCC 1303  
QY 1236 TGAAGAAATTATAGTTCTGTCAATAAGTACTGCTAGTTGAACGCTCTTCCGAAAGA 1295  
DB 1304 TGAAGAAATTATAGTTCTGTCAATAAGTACTGCTAGTTGAACGCTCTTCCGAAAGA 1363  
QY 1296 GTTCCAACCATTTTGAAGAAATGTTATGCTGTTATTGGATTAAAGTGTGATACCCAAT 1355  
DB 1364 GTTCCAACCATTTTGAAGAAATGTTATGCTGTTATTGGATTAAAGTGTGATACCCAAT 1423  
QY 1356 GAAATATCATTTCCAGAAAGTTGTCAAGTCTGTGATATTATATCATGATACAGCTCAAT 1415  
DB 1424 GAAATATCATTTCCAGAAAGTTGTCAAGTCTGTGATATTATATCATGATACAGCTCAAT 1483  
QY 1416 CTTGAAAACCTCCCTCTTAAACTTGAAGAAATGGAAGAGATGCTGTGTTCAATTGTCA 1475  
DB 1484 CTTGAAAACCTCCCTCTTAAACTTGAAGAAATGGAAGAGATGCTGTGTTCAATTGTCA 1543  
QY 1476 CTTTCCCTCATTTATCATTTGATGAATGTGATCAACCAACAAAGACAGTATATATAA 1535  
DB 1544 CTTTCCCTCATTTATCATTTGATGAATGTGATCAACCAACAAAGACAGTATATATAA 1603  
QY 1536 CATCATGAGGCAATATTGATGACAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAACAA 1595  
DB 1604 CATCATGAGGCAATATTGATGACAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAACAA 1663  
QY 1596 ACCAGTGAATCCCTCTTCTCAGATACCTGGGACTTAACAGCTTCACTCTGTGTTGAGGGGC 1655

```
Db 1664 ACCAGTGATTCCTCCCTCCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGC 1723
Qy 1656 CACGAAGCAAGCCAAAGCTGAAGACACACATTTTAAACTATGTGCCAATCTTGATGCAT 1715
Db 1724 CACGAAGCAAGCCAAAGCTGAAGACACACATTTTAAACTATGTGCCAATCTTGATGCAT 1783
Qy 1716 TACTATTAAACCTGTAAAGAAAACTTGATCACTGAAAAACCAATACAGAGCCATG 1775
Db 1784 TACTATTAAACCTGTAAAGAAAACTTGATCACTGAAAAACCAATACAGAGCCATG 1843
Qy 1776 CAAGAAGTTTGCCATTGCGAGATGCAACCAAGAGATCCATTTAAGAGAACTTCTAGA 1835
Db 1844 CAAGAAGTTTGCCATTGCGAGATGCAACCAAGAGATCCATTTAAGAGAACTTCTAGA 1903
Qy 1836 AATAATGACAAGGATTCAAACTTATGTCAAAATGAGTCCAATGTCAAGTTTGGAACTCA 1895
Db 1904 AATAATGACAAGGATTCAAACTTATGTCAAAATGAGTCCAATGTCAAGTTTGGAACTCA 1963
Qy 1896 ACCCTATGAACAATGGGCCATTCAAATGGAIAAAAAAGCTGCAAAAAAGAAATCGCAA 1955
Db 1964 ACCCTATGAACAATGGGCCATTCAAATGGAIAAAAAAGCTGCAAAAAAGAAATCGCAA 2023
Qy 1956 AGAACGTGTTGTGCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACAC 2015
Db 2024 AGAACGTGTTGTGCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACAC 2083
Qy 2016 AATTCGAATGATAGATGGCTACTCATCTTGAACTTTCTATAATGAGAGAAAGATAA 2075
Db 2084 AATTCGAATGATAGATGGCTACTCATCTTGAACTTTCTATAATGAGAGAAAGATAA 2143
Qy 2076 GAAGTTTGCACTCATAGAAAGATGATAGTGAAGGTGGTGAATGATGATTTGTGATGG 2135
Db 2144 GAAGTTTGCACTCATAGAAAGATGATAGTGAAGGTGGTGAATGATGATTTGTGATGG 2203
Qy 2136 TGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGATGAACAGATAGATTCT 2195
Db 2204 TGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGATGAACAGATAGATTCT 2263
Qy 2196 CATGACTTTATTTTGAACCAATAAATGTTGAAAAAGGCTGGCTGAAAAACCCAGATA 2255
Db 2264 CATGACTTTATTTTGAACCAATAAATGTTGAAAAAGGCTGGCTGAAAAACCCAGATA 2323
Qy 2256 TGAATAATGAAAAAGCTGACCAATTAAAGAAATACATTAATGAGCAATTAAGACTGA 2315
Db 2324 TGAATAATGAAAAAGCTGACCAATTAAAGAAATACATTAATGAGCAATTAAGACTGA 2383
Qy 2316 GGAATCAGCAGAGGAATATCTTTACAAAAACGACAGAGTGCATATGCGCTTCCCA 2375
Db 2384 GGAATCAGCAGAGGAATATCTTTACAAAAACGACAGAGTGCATATGCGCTTCCCA 2443
Qy 2376 GTGGAATTACTGAAAAATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCATCTGATTGG 2435
Db 2444 GTGGAATTACTGAAAAATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCATCTGATTGG 2503
Qy 2436 AGCTGACACAGCAGTGAAGTCAAAACCCATGACACAGATGAACAAAAAGAGTCATTAG 2495
Db 2504 AGCTGACACAGCAGTGAAGTCAAAACCCATGACACAGATGAACAAAAAGAGTCATTAG 2563
Qy 2496 TAAATTTGCGACTGGAATAATCAATCTGCTTATCGCTACACAGTGGCAGAGAAGTCT 2555
Db 2564 TAAATTTGCGACTGGAATAATCAATCTGCTTATCGCTACACAGTGGCAGAGAAGTCT 2623
Qy 2556 GGATATTAAAGATGTACATTTGTTATCCGTTATGCTCTGTCACCAATGAATAGCCAT 2615
Db 2624 GGATATTAAAGATGTACATTTGTTATCCGTTATGCTCTGTCACCAATGAATAGCCAT 2683
Qy 2616 GGTCCAGGCCCGGTGTCAGGCCAGAGCTGATGAGACACCTACGTCCTGGTGTCTCACAG 2675
Db 2684 GGTCCAGGCCCGGTGTCAGGCCAGAGCTGATGAGACACCTACGTCCTGGTGTCTCACAG 2743
Qy 2676 TGGTTCAGGAGTTATCGAATGAGACAGATTAATGATTTCCGAGAGAAAGATGATATAA 2735
Db 2744 TGGTTCAGGAGTTATCGAATGAGACAGATTAATGATTTCCGAGAGAAAGATGATATAA 2803
```

```
Qy 2736 AGCTATACATTGTGTTCAAAATATGAAACCAAGAGATATGCTCATTAAGATTTTGAATT 2795
Db 2804 AGCTATACATTGTGTTCAAAATATGAAACCAAGAGATATGCTCATTAAGATTTTGAATT 2863
Qy 2796 ACAGATGCAAAAGTAAATGGAIAAAAAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTA 2855
Db 2864 ACAGATGCAAAAGTAAATGGAIAAAAAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTA 2923
Qy 2856 CAAGATAACCCATCACTAATACTTCTTGCAAAAACTGCAAGTGTGCTAGCCTGTTTC 2915
Db 2924 CAAGATAACCCATCACTAATACTTCTTGCAAAAACTGCAAGTGTGCTAGCCTGTTTC 2983
Qy 2916 TGGGGAAGATATCATGTAATTGAGAAAAATGCATCAGCTCAATATGACCCAGAAATTCAA 2975
Db 2984 TGGGGAAGATATCATGTAATTGAGAAAAATGCATCAGCTCAATATGACCCAGAAATTCAA 3043
Qy 2976 GGAACCTTACATTGTAAAGAAAACAAGCACTGCAAAAAAGAGTGTGCCGACTATCAAAAT 3035
Db 3044 GGAACCTTACATTGTAAAGAAAACAAGCACTGCAAAAAAGAGTGTGCCGACTATCAAAAT 3103
Qy 3036 AATGCTGAATCATCTGCAAAATGTGGCCAGCCTTGGGGAACATGATGTCACAAAGG 3095
Db 3104 AATGCTGAATCATCTGCAAAATGTGGCCAGCCTTGGGGAACATGATGTCACAAAGG 3163
Qy 3096 CTTAGATTGCTGCTCTCAAAATTAAGAAATTTGTAGTGTGTTTCAAAATTAATTCAC 3155
Db 3164 CTTAGATTGCTGCTCTCAAAATTAAGAAATTTGTAGTGTGTTTCAAAATTAATTCAC 3223
Qy 3156 AAAGAAACAATACAAAAAGTGGGTAGAAATTAACCTATCACATTTCCCAATCTGACTATTC 3215
Db 3224 AAAGAAACAATACAAAAAGTGGGTAGAAATTAACCTATCACATTTCCCAATCTGACTATTC 3283
Qy 3216 AGAATGCTGTTTATTAAGTATGAGGATTAAGCACTTGATGAAGATTTCTTTAAATATCT 3275
Db 3284 AGAATGCTGTTTATTAAGTATGAGGATTAAGCACTTGATGAAGATTTCTTTAAATATCT 3343
Qy 3276 ATCAGTTAAACATTAATATGATTATGATTAATGTATTCTATTAAGTCTACAGAACTGACAT 3335
Db 3344 ATCAGTTAAACATTAATATGATTATGATTAATGTATTCTATTAAGTCTACAGAACTGACAT 3403
Qy 3336 AAGAATCAATTAATGATGTTGTTTACTCT 3364
Db 3404 AAGAATCAATTAATGATGTTTACTCT 3432

RESULT 13
AAD11170
ID AAD11170 standard; DNA; 3131 BP.
XX
AC AAD11170;
XX
DT 29-NOV-2001 (first entry)
XX
DE Human melanoma differentiation associated-5 protein-related DNA.
XX
KW Human melanoma differentiation associated gene; Mda-5; interferon; IFN;
KW RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
KW neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
KW breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
KW central nervous system; cytosstatic; apoptosis; ds.
XX
OS Homo sapiens.
XX
PN WO200164707-A1.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US006960.
XX
PR 29-FEB-2000; 2000US-00515363.
PA (UYCO ) UNIV COLUMBIA NEW YORK.
```



QY	1789	ATTGCAGATGCAACCCAGAGAAGATCCATTTTAAAGAGAACTTCTAGAAATTAATGACAAGG	1848
.Db	1618	ATTGCAGATGCAACCCAGAGAAGATCCATTTTAAAGAGAACTTCTAGAAATTAATGACAAGG	1677
QY	1849	ATTCAAACTTATTTGTCAAAAGAGTCCCATGTGCAGATTTTGGAACTCAACCCCTATGACAA	1908
Db	1678	ATTCAAACTTATTTGTCAAAAGAGTCCCATGTGCAGATTTTGGAACTCAACCCCTATGACAA	1737
QY	1909	TGGGCCATTCAAAATGGAAGAAAAAGCTGCAAAAAAGGAATCCGAAGAACGTGTTTGT	1968
Db	1738	TGGGCCATTCAAAATGGAAGAAAAAGCTGCAAAAAAGGAATCCGAAGAACGTGTTTGT	1797
QY	1969	GCAGAACATTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAATTCGAATGATA	2028
Db	1798	GCAGAACATTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAATTCGAATGATA	1857
QY	2029	GAT--GCGTATACTCATCTTGAAACTTTCTATTAATGAGAGAAAGATAAGAGTTTCAG	2086
Db	1858	GATCCGCGTATACTCATCTTGAAACTTTCTATTAATGAGAGAAAGATAAGAGTTTCAG	1917
QY	2087	TCATAGAAGATGATAGTGTGAGGGGTGTGATGATGATGATGATGATGATGATGATGATG	2146
Db	1918	TCATAGAAGATGATAGTGTGAGGGGTGTGATGATGATGATGATGATGATGATGATGATG	1977
QY	2147	AGGATGATTTAAGAAACCTTTGAAACTGTGATGAACAGATAGATTTCTCATGACTTTAT	2206
Db	1978	AGGATGATTTAAGAAACCTTTGAAACTGTGATGAACAGATAGATTTCTCATGACTTTAT	2037
QY	2207	TTTTGAAAAACAATAAAATGTTGAAAAAGGCTGCTGAAAAACCCAGATATGAAAAATGAAA	2266
Db	2038	TTTTGAAAAACAATAAAATGTTGAAAAAGGCTGCTGAAAAACCCAGATATGAAAAATGAAA	2097
QY	2267	AGCTGACCAATTAAGAAATACATAATGAGCAATTAATAAGACTGAGGAATCAGCAC	2326
Db	2098	AGCTGACCAATTAAGAAATACATAATGAGCAATTAATAAGACTGAGGAATCAGCAC	2157
QY	2327	GAGGAATTAATCTTTACAAAAACGACAGAGTGCATATGCGCTTCCAGTGGATTACTG	2386
Db	2158	GAGGAATTAATCTTTACAAAAACGACAGAGTGCATATGCGCTTCCAGTGGATTACTG	2217
QY	2387	AAAAAGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCACTGATTTGAGCTGACACA	2446
Db	2218	AAAAAGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCACTGATTTGAGCTGACACA	2277
QY	2447	GCAGTGAGTTCAAAACCCATGACACAGAATGAACAAAAAGAGTCATTAGTAATTTCCGA	2506
Db	2278	GCAGTGAGTTCAAAACCCATGACACAGAATGAACAAAAAGAGTCATTAGTAATTTCCGA	2337
QY	2507	CTGGAATAATCAATCTGCTTATCGCTACCAACAGTGGCAGAGAAAGGTCTGGATATTAAAG	2566
Db	2338	CTGGAATAATCAATCTGCTTATCGCTACCAACAGTGGCAGAGAAAGGTCTGGATATTAAAG	2397
QY	2567	AATGTAACATTTGTTATCCGTTATGCTCTCGTCAACCAATGAAATAGCCATGTCAGGCC	2626
Db	2398	AATGTAACATTTGTTATCCGTTATGCTCTCGTCAACCAATGAAATAGCCATGTCAGGCC	2457
QY	2627	GTCGTCGAGCCAGAGCTGATGAGACCACTACGTCCTGTGTTCTCACAGTGGTTCAGAG	2686
Db	2458	GTCGTCGAGCCAGAGCTGATGAGACCACTACGTCCTGTGTTCTCACAGTGGTTCAGAG	2517
QY	2687	TTATCGAACATGAGACAGTTAATGATTTCCGAGAGAAAGATGATGATTAAGCTATACATT	2746
Db	2518	TTATCGAACATGAGACAGTTAATGATTTCCGAGAGAAAGATGATGATTAAGCTATACATT	2577
QY	2747	GTCGTCGAAAAATATGAACCCAGAGAGTATGCTCATTAAGATTTTGGAAATTAAGATGCAAA	2806
Db	2578	GTCGTCGAAAAATATGAACCCAGAGAGTATGCTCATTAAGATTTTGGAAATTAAGATGCAAA	2637
QY	2807	GTAATAATGAAAAAGAAAAATGAACCAAGAGAAATATTGCCAAGCATTAACAAGATAACC	2866
Db	2638	GTAATAATGAAAAAGAAAAATGAACCAAGAGAAATATTGCCAAGCATTAACAAGATAACC	2697
QY	2867	CATCACTAATTAACCTTCCTTTGCAAAAACTGCAGTGTGCTAGCCCTGTTCTGGGGAGATA	2926

Accession	Sequence	Length (bp)	Organism	Feature
Db 2698	CATCACTAATACTTCCTTGGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATA	2757		
Qy 2927	TCCATGTAAATGAGAAAATGCATCAGTCAATATGACCCAGAAATTCAGGAATTACA	2986		
Db 2758	TCCATGTAAATGAGAAAATGCATCAGTCAATATGACCCAGAAATTCAGGAATTACA	2817		
Qy 2987	TTGTAAGAGAAAAACAAGCACTGCAGAAAAGAGTGTGCCGACTATCAAAATAAATG	3046		
Db 2818	TTGTAAGAGAAAAACAAGCACTGCAGAAAAGAGTGTGCCGACTATCAAAATAAATG	2877		
Qy 3047	TCATCTGCAAAATGTGGCCAGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGC	3106		
Db 2878	TCATCTGCAAAATGTGGCCAGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGC	2937		
Qy 3107	CTTGCTCTCAAAATAAGGAATTTGTAGTGTGTTCAAAAATAATTCAACAAGAAACAAT	3166		
Db 2938	CTTGCTCTCAAAATAAGGAATTTGTAGTGTGTTCAAAAATAATTCAACAAGAAACAAT	2997		
Qy 3167	ACAAAAAGTGGGTAGAAATTAACCTATCACATTTCCCAATCTTGACTATTGAGATGCTGTT	3226		
Db 2998	ACAAAAAGTGGGTAGAAATTAACCTATCACATTTCCCAATCTTGACTATTGAGATGCTGTT	3057		
Qy 3227	TATTTAGTAGAGGATTAACCTATGATTTGAAGATCTTTTAAATACTATCAGTTAAAC	3286		
Db 3058	TATTTAGTAGAGGATTAACCTATGATTTGAAGATCTTTTAAATACTATCAGTTAAAC	3117		
Qy 3287	ATTTAATATGATTA 3300			
Db 3118	ATTTAATATGATTA 3131			

RESULT 14  
ADJ75813 standard; DNA; 3771 BP.

Accession	Sequence	Length (bp)	Organism	Feature
ID ADJ75813	standard; DNA; 3771 BP.			
AC ADJ75813				
XX 20-MAY-2004	(first entry)			
XX Marker gene SEQ ID NO:1065.				
DE bronchial asthma; chronic obstructive pulmonary disease;				
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;				
KM gene therapy; marker gene; gene; ds.				
XX Mus musculus.				
OS Mus musculus.				
XX EPI394274-A2.				
PN 03-MAR-2004.				
XX 04-AUG-2003; 2003EP-00254857.				
PF 06-AUG-2002; 2002JP-00229312.				
PR 20-MAR-2003; 2003JP-00077212.				
XX (GENO-) GENOX RES INC.				
PA Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;				
PI WPI; 2004-193155/19.				
XX Testing for bronchial asthma or chronic obstructive pulmonary disease by				
PT comparing the expression level of a marker gene in a biological sample				
PT from a subject with the expression level of the gene in a sample from a				
PT healthy subject.				
XX Claim 14; SEQ ID NO 1065; 241pp; English.				
PS The present invention describes a method of testing for bronchial asthma				
CC or chronic obstructive pulmonary disease. The method comprises				
CC determining the expression level of a marker gene in a biological sample				



```
Db      1866 TGACACCAAGAAAATCCATTTAAAGAAACTTTAGAAATTATGGCAAGCATTCAGAC 1925
Oy      1857 TTAATGTCAAATGAGTCCAAATGTCAGATTTTGAAGCTCAACCCCTATGAACATGGCCAT 1916
Db      1926 TTACGCCCAAAAAAGTCCAATGTCAAGTTTGAACCCCAACATTATGAGCAGTGGCCAT 1985
Oy      1917 TCAATGCAAAAAAAGCTGCAAAAAAAGGAATCGCAAGAAGCTGTTTGCAGACA 1976
Db      1986 TCAATGAGAAAAAAGCTGCTAAAGACGGAATCGCAAGATCGCGCTGTGCGAGACA 2045
Oy      1977 TTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACAAATTCGAATGATGATCGCTA 2036
Db      2046 TTTCAGGAAGTACAACGAAGCCCTACAAATCAACGACACGATCCGAATGATGATGATA 2105
Oy      2037 TACTCATCTGAAACTTTCTAATGAAGAGAAAGATAAGATTGGCAGTCTAGAGA 2096
Db      2106 TAGCCACCTGGAGCATTCTACACTGATGAGAAAGAAAGATTCCGAGTCCT--CAA 2162
Oy      2097 TGAATGATGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2156
Db      2163 TGACAGCGACAAGAGTGAACGAGGCCAGCAGTTGCAATGACCACTTAAGGCCGATGT 2222
Oy      2157 AAAGAAACCTTGAACCTGGAACAGATAGATTCTCATGACTTTATTTTGA 2216
Db      2223 AAAGAAATCTTGAACCTGGAACAGATAGATTCTCATGATTTGTTTGA 2282
Oy      2217 CAATAAATGTTGAAAAGCTGGCTGAAAACCCAGAAATATGAAAAGCTGACCA 2276
Db      2283 CAAGAAATGTTGAAAAGCTGAAAACCCAAATATCGAAGATGAAAACCTCATTA 2342
Oy      2277 ATTAAGAAATACCATTAATGAGCAATATACTAGACTGAGGAATCAGCAGAGAAATAT 2336
Db      2343 ATTAGAAAACAGATACTGGAACAATTCACAAGGTGAGAGTCTCCCGAGGAATTAT 2402
Oy      2337 CTTTACAAAAACAGCAGAGTGCATATGCGCTTCCAGTGATTAAGTGA 2396
Db      2403 TTTTACAAAAACAGCAGAGACCTACGCACTTCCAGTGATCATGAAAAATGCA 2462
Oy      2397 ATTTGCTGAAGTAGAGTCAAAGCCCAACCATCTGATTGAGCTGACACAGAGTGA 2456
Db      2463 GTTTCGGGAAGTTGAGTCAAAGCCCATCACCTGATTGGCGCGGCACAGCAGTGA 2522
Oy      2457 CAAACCCATGACACAGAAATGAACAAAAGAGTCAATTAATTTCCGACTGGAAAAAT 2516
Db      2523 CAAGCCCATGACTCAGACTGAACAAAAGAGTCAATTAATTTCCGACTGGCGAAAT 2582
Oy      2517 CAATCTGCTTATCGCTACCAAGTGGCAGAGAGAGTCTGATATTAAAGATGTAACAT 2576
Db      2583 AAATCTGCTTATCGCTACGAGGTGGCAGAGAGGCTGGATATCAAGAGTGAATAT 2642
Oy      2577 TGTATCCGTTATGCTCTGTCACCAATGAATAGCCATGTCAGAGCCCGTGGTGAGC 2636
Db      2643 TGTATTCGTTATGCTTGTTCACGAACGAGATAGCCATGTCAGAGCCCGGGGTGAGC 2702
Oy      2637 CAGAGCTGATGAGAGCACCTACCTCTGCTGCTCAGAGTGTTCAGAGATTAATGACA 2696
Db      2703 CAGAGCTGATGAAAGCACGTAATGCTCTGCTCAGAGTGTTCAGAGATTAATGACA 2762
Oy      2697 TGAGACAGTAAATGATTTCCGAGAGAGATGATGATTAAGCTATACATTTGTTCAAAA 2756
Db      2763 GGAGATGTTAATGATTTCCGAGAGAGATGATGATTAAGCTATTAACCGTGTCAAAA 2822
Oy      2757 TATGAACACAGAGAGTATGCTCATTAAGATTTTGAATTAAGATGCAAGATTAATGA 2816
Db      2823 CATGAACACAGAGAGTATGCAATTAAGATTTTGAATTAAGATGCAAGATTAATGA 2882
Oy      2817 AAAGAAATGAAAAACAAGAAATATTGCAAGCATTAAGAATAACCACTCACTAAT 2876
Db      2883 AAAGAAATGAAAGTCAAAAGAACATTTGCAAGCAATTAACAAGCAATTCATGTTAAT 2942
Oy      2877 AACTTCTTTGCAAAACCTGAGTGTGCTAGCTGTTCTGGGAGAGATATCATGTAAT 2936
```

```
Db      2943 AACACTTCTGCAAAAAATTGTAGCATGCTGTGCTCGGAGAAAAACATCCATGTCAT 3002
Oy      2937 TGAGAAATATGCATCAGCTCAATATATGACCCCGAATTCAGGAACCTTAATTTGAAGAGA 2996
Db      3003 TGAGAAATATGCATCAGCTCAATATATGACACCCGAATTCAGGAACCTTAATTTGAAGAGA 3062
Oy      2997 AAACAAAGCACTGCAAAAGAGTGTGCCGACTATCAAAATTAATGTTGTTAAATCATCTGCA 3056
Db      3063 AAACAAAGCACTGCAAAAGAAATTTGCTGATTAATCAACCAATGAGAGATTAATCTGCA 3122
Oy      3057 ATGTGCCAGGCTTGGGAAACAATGATGTGCAAAAGGCTTAGATTGCTGTCTCA 3116
Db      3123 GTGTGCCAGGCTTGGGAAACAATGATGTGCAAAAGGCTTAGATTGCTGTCTCA 3182
Oy      3117 AATAAGAAATTTGTAGTGTTCAAAAATTAATCAACAAAGAAACAATACAAAAAAGTG 3176
Db      3183 AATAAGAAATTTGTAGTCAATTTCAAAATACTACCGAAGAAAGATTAAGAAGTG 3242
Oy      3177 GGTAGAATTACCTATCATATTTCCCAATCTTCACTATTCAGAATGCTGTTATTTAGTGA 3236
Db      3243 GGTGAATTTCCCTATCAGATTTCTGATCTTCACTACTCAGAAATACTGCTGTATAGTGA 3302
Oy      3237 TGAGGATTTAGCAGCTGATGGAAGATTTCTTTAAATACTATCATGTTAAACATTTAATATG 3296
Db      3303 TGAAGATTAGCAGCTGATCATGATTAATTAATACTGCCAACTCAACATTTGAATATG 3362
Oy      3297 ATTATGATTAATGATTCATTAATGCTACAGAACTGACATAAGAATCAATTA 3348
Db      3363 ATTTAATTTGTTTTCGCTACACTACTGAGCTAAGCTAAGATATGTA 3414
```

```
RESULT 15
AAS40960
ID AAS40960 standard; cDNA; 1967 BP.
XX AC AAS40960;
XX AC
DT 17-DEC-2001 (first entry)
XX DE
DE cDNA encoding novel human enzyme polypeptide #176.
XX XX
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
XX OS Homo sapiens.
XX PN WO200155301-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001239.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
```

PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 25-SEP-2000; 2000US-0234999P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001.465566/50.  
DR P-PSDB; AAU23090.  
XX  
XX  
PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,  
PT treating neural, immune system, muscular, reproductive, pulmonary,  
PT cardiovascular, renal, proliferative disorders and cancerous diseases.  
XX  
XX  
PS Claim 4; SEQ ID NO 186; 1180pp; English.  
XX  
CC The present invention relates to the isolation of novel human enzyme  
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences  
CC encoding them. The enzyme polypeptides of the invention may comprise the  
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,  
CC isomerases or ligases. The sequences of the invention are useful in the  
CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
CC disorders including hyperproliferative disorders (e.g. cancer),  
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.  
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic  
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),  
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders  
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and  
CC infectious disorders (e.g. Influenza). The polynucleotides of the

CC invention can also be used in gene therapy. AAS40785-AAS41684 represent  
CC cDNA sequences encoding for the novel human enzyme polypeptides of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1967 BP; 723 A; 344 C; 416 G; 476 T; 0 U; 8 Other;

Query Match 55.8%; Score 1878; DB 4; Length 1967;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 1941; Conservative 5; Mismatches 18; Indels 7; Gaps 5;

QY 1047 AGAGAATGTGCAGCAAGACATCCCGAGCCGAACTCCAGCTCAGGCTTACCAAT 1106  
DB 2 AGAGAATGTGCAGCAAGACATCCCGAGCCGAACTCCAGCTCAGGCTTACCAAT 61  
QY 1107 GGAAGTTGCCCGCAGCCTTGGAGGGAAGATATCATCTGCTCCCTACAGGAG 1166  
DB 62 GGAAGTTGCCCGCAGCCTTGGAGGGAAGATATCATCTGCTCCCTACAGGAG 121  
QY 1167 TGGAAAAAACAGAGTGGCTGTTTACATTGCCAAGATCACTTAGACAAGAAAAAGC 1226  
DB 122 TGGAAAAAACAGAGTGGCTGTTTACATTGCCAAGATCACTTAGACAAGAAAAAGC 181  
QY 1227 ATCTGAGCCTGAAAAAGTTATAGTCTGTGCAATAGTAAGTCTAGTTGAACAGCTCTT 1286  
DB 182 ATCTGAGCCTGAAAAAGTTATAGTCTGTGCAATAGTAAGTCTAGTTGAACAGCTCTT 241  
QY 1287 CCGCAAGAGGTTCCAAACATTTTGAAGAAATGATCGTATTGATTAAGTGGA 1346  
DB 242 CCGCAAGAGGTTCCAAACATTTTGAAGAAATGATCGTATTGATTAAGTGGA 301  
QY 1347 TACCCAAGTGAATATATATTTCCAGAGAGTTGTCAGTCTCTGATATATATCAGTAC 1406  
DB 302 TACCCAAGTGAATATATATTTCCAGAGAGTTGTCAGTCTCTGATATATATCAGTAC 361  
QY 1407 AGCTCAATCTTGAAGAACTCCCTCTTAACCTTGAAAAATGAGAGAGTGTGTCTCA 1466  
DB 362 AGCTCAATCTTGAAGAACTCCCTCTTAACCTTGAAAAATGAGAGAGTGTGTCTCA 421  
QY 1467 ATGTGACAGCTTTTCCCTCATTTATCATTTGATGAATGTATCACAACCAAGAGCAGT 1526  
DB 422 ATGTGACAGCTTTTCCCTCATTTATCATTTGATGAATGTATCACAACCAAGAGCAGT 481  
QY 1527 GTATATAATACATCATGAGGATTTATTTGATGACAGAGTTGAAAAACAATAGACTCAAGAA 1586  
DB 482 GTATATAATACATCATGAGGATTTATTTGATGACAGAGTTGAAAAACAATAGACTCAAGAA 541  
QY 1587 AGAAAAACAACAGTATTCCTCTCTCAGATACCTGGACTAACAGCTTCACTGTGT 1646  
DB 542 AGAAAAACAACAGTATTCCTCTCTCAGATACCTGGACTAACAGCTTCACTGTGT 601  
QY 1647 TGGAGGGGGCCAGAGCAAGCCAAAGCTGAAGAACATTTTAAACTATGTGCCAATCT 1706  
DB 602 TGGAGGGGGCCAGAGCAAGCCAAAGCTGAAGAACATTTTAAACTATGTGCCAATCT 661  
QY 1707 TGAATGATTTACTATTAATACTGTTAAAGAAAACTTGATCAACTGAAGAAAAACAATACA 1766  
DB 662 TGAATGATTTACTATTAATACTGTTAAAGAAAACTTGATCAACTGAAGAAAAACAATACA 721  
QY 1767 GGAGCCATGCAAGAGTTGCCATTTGAGATGCAACCAAGAGATCCATTTAAAGAGAA 1826  
DB 722 GGAGCCATGCAAGAGTTGCCATTTGAGATGCAACCAAGAGATCCATTTAAAGAGAA 781  
QY 1827 ACTTCTAGAAATATGACAAGATTCAAACTTATTTGTCAAATGAGTCCATGTCAAGTTT 1886  
DB 782 ACTTCTAGAAATATGACAAGATTCAAACTTATTTGTCAAATGAGTCCATGTCAAGTTT 841  
QY 1887 TGGAACTCAACCTTATGAACAATGGGCATTCAAATGAAAAAAAAGCTGCAAAAAAGG 1946  
DB 842 TGGAACTCAACCTTATGAACAATGGGCATTCAAATGAAAAAAAAGCTGCAAAAAAGG 901  
QY 1947 AAATCGCAAAAGACGTGTTTGTGACAGAACATTTGAGGAAGTACAAATGAGCCCTTCAAAAT 2006

DB 902 AAATCGCAAAAGACGTGTTTGTGACAGAACATTTGAGGAAGTACAAATGAGCCCTTACAAAT 961  
QY 2007 TAATGACACAATTCGATGATAGATGCGTATATCTCATCTTGAACCTTCTATATGAAGA 2066  
DB 962 TAATGACACAATTCGATGATAGATGCGTATATCTCATCTTGAACCTTCTATATGAAGA 1021  
QY 2067 GAAAGATAGAAGTTTGCAGTCATAGAGATAGATGATGAGGCTGATGATGAGTA 2126  
DB 1022 GAAAGATAGAAGTTTGCAGTCATAGAGATAGATGATGAGGCTGATGATGAGTA 1081  
QY 2127 TTGTGATGATGATGAGATGAGATGATTTTAAAGAAACCTTTGAACCTGATGAACAGA 2186  
DB 1082 TTGTGATGATGATGAGATGAGATGATTTTAAAGAAACCTTTGAACCTGATGAACAGA 1141  
QY 2187 TAGATTTCTCATGACTTTATTTTGAAGAAACAATAAATGTTGAAGAGCTGCTGAAAA 2246  
DB 1142 TAGATTTCTCATGACTTTATTTTGAAGAAACAATAAATGTTGAAGAGCTGCTGAAAA 1201  
QY 2247 CCCAGATATGAAAAATGAAAAAGCTGACCAAAATTAAGAAATACCAATATGAGCAATATAC 2306  
DB 1202 CCCAGATATGAAAAATGAAAAAGCTGACCAAAATTAAGAAATACCAATATGAGCAATATAC 1261  
QY 2307 TAGGACTAGGAATCAGACAGAGGATATCTTTACAAAAACAGACAGAGTGCATATGC 2366  
DB 1262 TAGGACTAGGAATCAGACAGAGGATATCTTTACAAAAACAGACAGAGTGCATATGC 1321  
QY 2367 GCTTTCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCAACA 2426  
DB 1322 GCTTTCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCAACA 1381  
QY 2427 TCTGATTTGAGCTGACACAGCAGTGAAGTTCAAAACCATGACACAGATGAACAAAAAGA 2486  
DB 1382 TCTGATTTGAGCTGACACAGCAGTGAAGTTCAAAACCATGACACAGATGAACAAAAAGA 1441  
QY 2487 AGTCATTAGTAATTTGCACTGAAAAATCAATCTGTTATCGTTACCAAGTGACAGA 2546  
DB 1442 AGTCATTAGTAATTTGCACTGAAAAATCAATCTGTTATCGCT-CCACAGTGACAGA 1500  
QY 2547 AGAAGTCTGATTAATAAGATGTAACATTTGTTATCCGTTATGCTCTCCACCAATGA 2606  
DB 1501 AGAAGTCTGATTAATAAGATGTAACATTTGTTATCCGTTATGCTCTCCACCAATGA 1560  
QY 2607 AATAGCATGTGCCAGGCCCGTGTGAGCCAGAGCTGATGAGAGCACTACGCTGTGT 2666  
DB 1561 AATA-CCATGTGTCA-GCCCGTGTGAGCTAGAGCTGATGAGAGCACTACGCT-CTGTGT 1617  
QY 2667 TGCTCACAGTGTTCAGAGGTTATGGAACATGAGACAGTGAATGATTTCCGAGAGAGAT 2726  
DB 1618 TGCTCACAGTGTTCAGAGGTTATGGAACATGAGACAGTGAATGATTTCCGAGAGAGAT 1677  
QY 2727 GATGATTAAGCTATACATTTGTTCAAAATATGAAGACAGAGGATATGCTCATAGAT 2786  
DB 1678 GATGATTAAGCTATACATTTGTTCAAAATATGAAGACAGAGGATATGCTCATAGAT 1734  
QY 2787 TTTGGAATTACAGATGCAAGTATATATGAAAAAGAAATGAAACCAAGAGAAATATTGC 2846  
DB 1735 TTTGGAATTACAGATGCMAGTATATATGAAAAAGAAATGAAACCAAGAGAAATATTGC 1794  
QY 2847 CAAGCATTAAGAAATTAACCCATCACTAATACTTTCTTTGCAAAAACTGCAGTGTGT 2906  
DB 1795 CAAGCATTAAGAAATTAACCCATCACTAATACTTTCTTTGCAAAAACTGCAGTGTGT 1854  
QY 2907 AGCCTGTTCTGGGAGATATCCATGTAATGAGAAAAATGATCAAGTCAATATGACCCC 2966  
DB 1855 AGCCTGTTCTGGGAGATATCCATGTAATGAGAAAAATGATCAATATGACCCC 1914  
QY 2967 AGAATTCAAGGAACCTTACATTTGTAAGAGAAAACAAAGCACTGCAAAAAAGAA 3017  
DB 1915 AGAATTCAAGGAACCTTACATTTGTAAGAGAAAACAAAGCACTGCAAAAAAGAA 1965

Job time : 1898 secs

---

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2006, 18:48:11 ; Search time 11783 Seconds  
(without alignments)  
13361.498 Million cell updates/sec

Title: US-09-515-363C-1  
Perfect score: 3365  
Sequence: 1 gcgcgcgcgcgtgagagccc.....aaaatgattgttctctg 3365

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1:	gb_est1:*
2:	gb_est2:*
3:	gb_est3:*
4:	gb_htc:*
5:	gb_est4:*
6:	gb_est5:*
7:	gb_est6:*
8:	gb_est7:*
9:	gb_gss1:*
10:	gb_gss2:*
11:	gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3052.6	90.7	3078	11	DQ050960 Homo sapi
2	2861	85.0	3078	11	DQ050961 Pan trogl
3	1393	41.4	2304	4	AK037057 Mus muscu
4	1367.4	40.6	2331	4	AK040519 Mus muscu
5	926.4	27.5	1013	3	BQ233683 AGENCOURT
6	900.4	26.8	1174	5	BU902097 AGENCOURT
7	806.8	24.0	1115	3	BM467983 AGENCOURT
8	801.8	23.8	870	5	BQ960157 AGENCOURT
9	728.2	21.6	918	5	BU189982 AGENCOURT
10	718.8	21.4	755	5	BQ772836 UI-H-FEO-
11	711.6	21.1	1035	2	BF337464 602035195
12	687	20.4	729	2	BF983236 602305873
13	680.4	20.2	690	8	DR422160 nav08g05.
14	654.4	19.4	701	6	CA423868 UI-H-FEI-
15	652.8	19.4	827	2	BG741146 602631817
16	651.8	19.4	781	2	BF686405 602143786
17	641.8	19.1	653	7	CR752055 DKFZp469M
18	626	18.6	1046	3	BM476961 AGENCOURT
19	625.4	18.6	627	5	BX492926 DKFZp781C
20	622.8	18.5	1041	2	BE882040 601505326
21	618.4	18.4	1197	4	AK018602 Mus muscu
22	616.2	18.3	1239	3	BM467774 AGENCOURT

23	606.4	18.0	972	5	BY720783	BY720783
24	599.2	17.8	672	3	BQ316075	BQ316075 CM3-CT027
25	599.2	17.8	672	3	BQ316108	BQ316108 CM3-CT027
26	599.2	17.8	672	3	BQ316120	BQ316120 CM3-CT027
27	596.8	17.7	619	3	BQ016388	BQ016388 UI-H-DT1-
28	584	17.4	1611	3	BM455238	BM455238 AGENCOURT
29	580.4	17.2	582	1	AU310289	AU310289 AU310289
30	559.8	16.6	581	3	BP234383	BP234383 BP234383
31	551.2	16.4	568	3	BQ311714	BQ311714 QV3-BN004
32	549.8	16.3	579	5	BU685900	BU685900 UI-CF-DU1
33	547.8	16.3	671	6	CB453859	CB453859 709752 MA
34	544	16.2	544	1	AU310290	AU310290 AU310290
35	538	16.0	585	3	BP210040	BP210040 BP210040
36	533.8	15.9	674	7	CV194473	CV194473 SNESTbab0
37	531.4	15.8	536	1	AI718277	AI718277 as51e06.x
38	525.4	15.6	647	6	CB424267	CB424267 598511 MA
39	523.6	15.6	698	7	CO748417	CO748417 SNESTbaa6
40	522.4	15.5	1201	3	BM551088	BM551088 AGENCOURT
41	522.2	15.5	574	3	BP297266	BP297266 BP297266
42	521	15.5	682	6	CF363618	CF363618 833509 MA
43	520.4	15.5	537	1	AI806204	AI806204 wF26C12.X
44	517.8	15.4	800	7	CO396840	CO396840 AGENCOURT
45	517	15.4	680	6	CF365022	CF365022 835621 MA

ALIGNMENTS

RESULT 1	DQ050960	3078 bp	DNA	linear	GSS 02-JUN-2005
LOCUS	DQ050960				
DEFINITION	Homo sapiens MDAS gene, VIRTUAL TRANSCRIPT, partial sequence.				
ACCESSION	DQ050960				
VERSION	DQ050960.1	GI:66904159			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3078)				
AUTHORS	Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeille,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees				
JOURNAL	(ex) Plos Biol. 3 (6), E170 (2005)				
PUBMED	15869325				
REFERENCE	2 (bases 1 to 3078)				
AUTHORS	Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeille,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				
COMMENT	Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
FEATURES	This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.				
source	location/Qualifiers				
	1..3078				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="2"				
	<1..>3078				
	/gene="MDAS"				
	/locus_tag="HC18981"				
ORIGIN					
Query Match	90.7%;	Score 3052.6;	DB 11;	Length 3078;	
Best Local Similarity	99.3%;	Pred. No. 0;			
Matches 3055;	Conservative 0;	Mismatches 23;	Indels 0;	Gaps 0;	

Qy	169	ATGTCGAATGGGTATTTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGCCTTCAGGGCC	228
Db	1	ATGTGCAATGGGTATTTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGCCTTCAGGGCC	60
Qy	229	AGGGTAAAAATGTATCATCTCCAGGTGGAGCCCTGTGCTGGACTAACCTGACCTTTCTGCCTGCA	288
Db	61	AGGGTAAAAATGTATCATCTCCAGGTGGAGCCCTGTGCTGGACTAACCTGACCTTTCTGCCTGCA	120
Qy	289	GAGGTGAAGAGCAGATTCAAGAGCAGTCCGCACTCCGGAACATGCAGGCAAGTTGAA	348
Db	121	GAGGTGAAGAGCAGATTCAAGAGCAGTCCGCACTCCGGAACATGCAGGCAAGTTGAA	180
Qy	349	CTGCTGCTGAGCACCTTTGGAAGAGGAGTCTGGCACCTTGGTTGGACTCCGGAAATTCGTG	408
Db	181	CTGCTGCTGAGCACCTTTGGAAGAGGAGTCTGGCACCTTGGTTGGACTCCGGAAATTCGTG	240
Qy	409	GAGGCCCTCCGAGAACCGGCAGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCACG	468
Db	241	GAGGCCCTCCGAGAACCGGCAGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCACG	300
Qy	469	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCACTGCTGAACCTC	528
Db	301	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCACTGCTGAACCTC	360
Qy	529	CTTCAGCCCACTCTGGTGCACAAGCTTCTAGTTAGAACGTCTTGATAAGTGCATGGAG	588
Db	361	CTTCAGCCCACTCTGGTGCACAAGCTTCTAGTTAGAACGTCTTGATAAGTGCATGGAG	420
Qy	589	GAGGAACCTGTTGACAATTGAAGAACAAGAACCGGATTGCTCTGCAGAAAACAATGGAAT	648
Db	421	GAGGAACCTGTTGACAATTGAAGAACAAGAACCGGATTGCTCTGCAGAAAACAATGGAAT	480
Qy	649	GAATCAGGTGTAAAGAGCTACTAAAAAGGATTGTGAGAAAGAAAACTGGTTCTCTGCA	708
Db	481	GAATCAGGTGTAAAGAGCTACTAAAAAGGATTGTGAGAAAGAAAACTGGTTCTCTGCA	540
Qy	709	TTTCTGAATGTTCTTCGTCAAACAGGAAACAATGAATCTGTCCAAGATTAAACAGCTCT	768
Db	541	TTTCTGAATGTTCTTCGTCAAACAGGAAACAATGAATCTGTCCAAGATTAAACAGCTCT	600
Qy	769	GATTGCTCAGAAAGCAATGCAGAGATTGGAATTTATCACAAGTTGATGTCCTCAAGTG	828
Db	601	GATTGCTCAGAAAGCAATGCAGAGATTGGAATTTATCACAAGTTGATGTCCTCAAGTG	660
Qy	829	GAAAGCAACTTCTTTCAACCAACAGTTCAGCCAAATCTGGAGAAAGAGTCTGGGGCATG	888
Db	661	GAAAGCAACTTCTTTCAACCAACAGTTCAGCCAAATCTGGAGAAAGAGTCTGGGGCATG	720
Qy	889	GAGAATACTCATCAGAAATCATCTTTTGAGATTCTTCTGTAGTTTCAGAAATCAGACACA	948
Db	721	GAGAATACTCATCAGAAATCATCTTTTGAGATTCTTCTGTAGTTTCAGAAATCAGACACA	780
Qy	949	AGTTTGCAGAAAGAAAGTCTCAGCTGCTTAGATGAAGTCTTGACATAACAGACATG	1008
Db	781	AGTTTGCAGAAAGAAAGTCTCAGCTGCTTAGATGAAGTCTTGACATAACAGACATG	840
Qy	1009	GGCAGTGATTCAGGCAACCATGGGAAAGTGATTCAAGATGAAGAAATGTGGCAGCAAGACA	1068
Db	841	GGCAGTGATTCAGGCAACCATGGGAAAGTGATTCAAGATGAAGAAATGTGGCAGCAAGACA	900
Qy	1069	TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAAATGGAAGTTGCCACGCCAGCTTG	1128
Db	901	TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAAATGGAAGTTGCCACGCCAGCTTG	960
Qy	1129	GAAAGGAAGATATCATCTCTGCTCCCTACAGGGAAGTGAAGAAAAACAGAGTGGCTGTT	1188
Db	961	GAAAGGAAGATATCATCTCTGCTCCCTACAGGGAAGTGAAGAAAAACAGAGTGGCTGTT	1020
Qy	1189	TACATTGCCAAGATCACTTAGACAAGAAGAAAAAGCATCTGAGCCTGGAAGATTATA	1248
Db	1021	TACATTGCCAAGATCACTTAGACAAGAAGAAAAAGCATCTGAGCCTGGAAGATTATA	1080

[illegible]

Db 2161 GGAATTAATCTTTACAAAAACGACAGAGTGCAATGCGCTTCCAGTGAATTAAGAA 2220  
Oy 2389 AATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCACTGATTGAGCTGACACAGC 2448  
Db 2221 AATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCACTGATTGAGCTGACACAGC 2280  
Oy 2449 AGTGAAGTTCAAAACCCATGACACAGAATGAACAAAAAGAGTCAATAGTAAATTTCCACT 2508  
Db 2281 AGTGAAGTTCAAAACCCATGACACAGAATGAACAAAAAGAGTCAATAGTAAATTTCCACT 2340  
Oy 2509 GGAATAATCAATCTGCTTATCGCTACACAGTGGCAGAAAGGTCTGATATTAAGAA 2568  
Db 2341 GGAATAATCAATCTGCTTATCGCTACACAGTGGCAGAAAGGTCTGATATTAAGAA 2400  
Oy 2569 TGTACATTTGTTATCCGTTATGCTCTGTCACCAATGAATAAGCCATGTCAGGCCCGT 2628  
Db 2401 TGTACATTTGTTATCCGTTATGCTCTGTCACCAATGAATAAGCCATGTCAGGCCCGT 2460  
Oy 2629 GGTGAGCCAGAGCTGATGAGAGCACTACGTCCTGTTGCTCACAGTGGTTCAGAGTT 2688  
Db 2461 GGTGAGCCAGAGCTGATGAGAGCACTACGTCCTGTTGCTCACAGTGGTTCAGAGTT 2520  
Oy 2689 ATCGAATGAGACAGTTAATGATTTCCGAGAGAAAGATGATATAAGCTATACATGTT 2748  
Db 2521 ATCGAATGAGACAGTTAATGATTTCCGAGAGAAAGATGATATAAGCTATACATGTT 2580  
Oy 2749 GTTCAAAATATGAACACAGAGAGTATGCTCATTAAGATTTTGAATTCAGATGCAAAAGT 2808  
Db 2581 GTTCAAAATATGAACACAGAGAGTATGCTCATTAAGATTTTGAATTCAGATGCAAAAGT 2640  
Oy 2809 ATTAATGAAAAAGAAAAATGAACCAAGAAATATTGCCAAGCATTACAAGATTAACCCA 2868  
Db 2641 ATTAATGAAAAAGAAAAATGAACCAAGAAATATTGCCAAGCATTACAAGATTAACCCA 2700  
Oy 2869 TCACATAATACTTTCTTGCAAAACTGCAGTGTGCTAGCCTGTTCTGGGAAAGATATC 2928  
Db 2701 TCACATAATACTTTCTTGCAAAACTGCAGTGTGCTAGCCTGTTCTGGGAAAGATATC 2760  
Oy 2929 CATGTAATTGAGAAAAATGCATCACTCATATGACCCCAAGATTCAAGGAACTTTACATT 2988  
Db 2761 CATGTAATTGAGAAAAATGCATCACTCATATGACCCCAAGATTCAAGGAACTTTACATT 2820  
Oy 2989 GTTAAGAAAAACAAAGCACTGCAAAAAGAGTGTGCCGACTATCAATTAATGTGAAATC 3048  
Db 2821 GTTAAGAAAAACAAAGCACTGCAAAAAGAGTGTGCCGACTATCAATTAATGTGAAATC 2880  
Oy 3049 ATCTGCAAAATGTGGCCAGGCTTGGGAAACAATGATGTGCACAAAGGCTTAGATTGCT 3108  
Db 2881 ATCTGCAAAATGTGGCCAGGCTTGGGAAACAATGATGTGCACAAAGGCTTAGATTGCT 2940  
Oy 3109 TGTCTCAAAATAAGGAATTTGTAGTGTGTTTCAAAAAATAATTCACAAAGAAACAATAC 3168  
Db 2941 TGTCTCAAAATAAGGAATTTGTAGTGTGTTTCAAAAAATAATTCACAAAGAAACAATAC 3000  
Oy 3169 AAAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAGAATGCTGTTA 3228  
Db 3001 AAAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAGAATGCTGTTA 3060  
Oy 3229 TTTAGTATGAGGATTAG 3246  
Db 3061 TTTAGTATGAGGATTAG 3078

RESULT 2  
DQ050961 3078 bp DNA linear GSS 02-JUN-2005  
LOCUS DQ050961 Pan troglodytes MDAs gene, VIRUAL TRANSCRIPT, partial sequence,  
DEFINITION genomic survey sequence.  
ACCESSION DQ050961  
VERSION DQ050961.1 GI:66904160  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Pan.  
REFERENCE 1 (bases 1 to 3078)  
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeille,D.,  
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and  
Chimpanzees  
JOURNAL (er) Plos Biol. 3 (6), E170 (2005)  
PUBMED 15869325  
REFERENCE 2 (bases 1 to 3078)  
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeille,D.,  
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment. Translation starts at the beginning of  
alignment.  
FEATURES  
source location/Qualifiers  
1..3078  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>3078  
/gene="MDA5"  
/locus\_tag="HC18981"  
ORIGIN  
Query Match 85.0%; Score 2861; DB 11; Length 3078;  
Best Local Similarity 93.1%; Pred. No. 0;  
Matches 2867; Conservative 0; Mismatches 211; Indels 0; Gaps 0;  
Oy 169 ATGTGCAATGGGTATTCCACAGACGAGAAATTCGGCTATCTCATCTCGTGTTCAGGGCC 228  
Db 1 ATGTGCAATGGGTATTCCACAGACGAGAAATTCGGCTATCTCATCTCGTGTTCAGGGCC 60  
Oy 229 AGGCTGAAAAATGTACATCCAGGTGAGCCGTGTGCTGACTACCTGACCTTTCTGCCTGCA 288  
Db 61 AGGCTGAAAAATGTACATCCAGGTGAGCCGTGTGCTGACTACCTGACCTTTCTGCCTGCA 120  
Oy 289 GAGTGAAGAGCAGATTGAGAGACAGTGGCACTCCGGGAACATGACGCACTTGA 348  
Db 121 GAGTGAAGAGCAGATTGAGAGACAGTGGCACTCCGGGAACATGACGCACTTGA 180  
Oy 349 CTGCTGCTGAGCACCTTGAGAAAGGAGTCTGGCACCTTGTGACTCGGGAATTGCTG 408  
Db 181 CTGCTGCTGAGCACCTTGAGAAAGGAGTCTGGCACCTTGTGACTCGGGAATTGCTG 240  
Oy 409 GAGGCCCTCCGAGAACCGGAGCCCTCTGGCCCGGCTACATGAACCTGAGCTCAGC 468  
Db 241 GAGGCCCTCCGAGAACCGGAGCCCTCTGGCCCGGCTACATGAACCTGAGCTCAGC 300  
Oy 469 GACTTGCCTCTCCATCGTTTGAGAAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 528  
Db 301 GACTTGCCTCTCCATCGTTTGAGAAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 360  
Oy 529 CTTGAGCCCACTCTGTGGAACAAGCTTCTAGTTAGAGACGTCCTTGATTAAGTGCATGAG 588  
Db 361 CTTGAGCCCACTCTGTGGAACAAGCTTCTAGTTAGAGACGTCCTTGATTAAGTGCATGAG 420  
Oy 589 GAGGAAGTGTGACAAATTGAAGACAGAAACCGGATTGCTGCTGAGAAAACAATGGAAT 648  
Db 421 GAGGAAGTGTGACAAATTGAAGACAGAAACCGGATTGCTGCTGAGAAAACAATGGAAT 480  
Oy 649 GAATCAGGTGTAAGAGAGCTATAAAAAGGATTGTGCAAGAAAGAAACTGTTCTCTGCA 708  
Db 481 GAGTCAGGTGTAAGAGAGCTATAAAAAGGATTGTGCAAGAAAGAAACTGTTCTCTGCA 540  
Oy 709 TTTCTGAATGTTCTTGTGTAACAAGAAACATGAACCTTGTCCAAGAGTTAAAGGCTCT 768

Db	541	TTTCTGATGTTCTTCGTCAAACAGGAAACAATGACTGTCCAGAGTTAACAGGCTCT	600
Qy	769	GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTATCACAAGTTGATGGTCTCAAGTG	828
Db	601	GATTGCTCAGAAAGCAATGCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	660
Qy	829	GAAGAGCAACTTCTTCAACCAACAGTTCAGCCAAATCTGGAGAAAGAGGTCTGGGGCATG	888
Db	661	NN	720
Qy	889	GAGATAACTCATCAGAAATCATCTTTGCAGATTCTTCTGAGTTTCAGAAATCAGACACA	948
Db	721	NN	780
Qy	949	AGTTGGCAGAAAGAAAGTGTACGCTGCTTAGATGAAGTCTTGACATTAACAGCAACATG	1008
Db	781	AGTTGGCAGAAAGAAAGTGTACGCTGCTTAGATGAAGTCTTGACATTAACAGCAACATG	840
Qy	1009	GGCAGTGAATCAGGCACCATGGGAAGTGAATCAGATGAAGAGAAATGTGGCAGCAAGACA	1068
Db	841	GGCAGTGAATCAGGCACCATGGGAAGTGAATCAGATGAAGAGAAATGTGGCAGCAAGACA	900
Qy	1069	TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAAGTTGCCAGCCAGCTTG	1128
Db	901	TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAAGTTGCCAGCCAGCTTG	960
Qy	1129	GAAAGGAAGATATCATCTCTGCTCCCTACAGGAGTGGAAAAAACAGAGTGGCTGTT	1188
Db	961	GAAAGGAAGATATCATCTCTGCTCCCTACAGGAGTGGAAAAAACAGAGTGGCTGTT	1020
Qy	1189	TACATTGCCAAGGATCACTTAGACAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATA	1248
Db	1021	TACATTGCCAAGGATCACTTAGACAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATA	1080
Qy	1249	GTTCTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCGCAAGAGTTCCAAACATTT	1308
Db	1081	GTTCTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCGCAAGAGTTCCAAACATTT	1140
Qy	1309	TTGAAGAAATGTTATCGTGTTATTGGATTAAAGTGATACCCAACTGAAAAATATCATTT	1368
Db	1141	TTGAAGAAATGTTATCGTGTTATTGGATTAAAGTGATACCCAACTGAAAAATATCATTT	1200
Qy	1369	CCAGAAGTTGTCAAGTCTGTGATATTATATCAGTACAGCTCAAACTCCTGAAAACTCC	1428
Db	1201	CCAGAAGTTGTCAAGTCTGTGATATTATATCAGTACAGCTCAAACTCCTGAAAACTCC	1260
Qy	1429	CTCTTAACTTGAAAAATGGAAGAAGTGTGTTTCAATTGTCAGACTTTTCCCTCATTT	1488
Db	1261	CTCTTAACTTGAAAAATGGAAGAAGTGTGTTTCAATTGTCAGACTTTTCCCTCATTT	1320
Qy	1489	ATCATTTGATGATGTGTATCACCAACAAGAAAGCAGTGTAATAACATCATGAGGCAT	1548
Db	1321	ATCATTTGATGATGTGTATCACCAACAAGAAAGCAGTGTAATAACATCATGAGGCAT	1380
Qy	1549	TATTTGATGAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAACAAACAGTGAATCCC	1608
Db	1381	TATTTGATGAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAACAAACAGTGAATCCC	1440
Qy	1609	CTTCCCTCAGATACTGGGACTAAACAGCTTCACTGGTGTGGAGGGGCCAGAAAGCAAGCC	1668
Db	1441	CTTCCCTCAGATACTGGGACTAAACAGCTTCACTGGTGTGGAGGGGCCAGAAAGCAAGCC	1500
Qy	1669	AAAGCTGAAGAACACATTTTAAACTATATGTGCCAATCTTGATGCATTTACTATTAAACT	1728
Db	1501	AAAGCTGAAGAACACATTTTAAACTATATGTGCCAATCTTGATGCATTTACTATTAAACT	1560
Qy	1729	GTTAAAGAAAACTTTGATCACTGAAAAAACCAATATACAGAGCCATGCAAGAAAGTTGCC	1788
Db	1561	GTTAAAGAAAACTTTGATCACTGAAAAAACCAATATACAGAGCCATGCAAGAAAGTTGCC	1620
Qy	1789	ATTGAGATGCAACCAAGAAAGATCCATTTAAAGAAACTTCTAGAATTAATGACAAG	1848

Db	1621	ATTGCAGATGCAACCAGAGAAAGATCCATTTTAAAGAGAAACTNNMAGAAATATATGACAAGG	1680
QY	1849	ATTCAAACCTTAATTGTCAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCCTATGAAACA	1908
Db	1681	ATTCAACCTTAATTGTCAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCCTATGAAACA	1740
QY	1909	TGGGCCATTTCAAATGGAAGAAAAAGCTGCAGAAAAAGGAAATCGCAAGAACGTGTTGT	1968
Db	1741	TGGGCCATTTCAAATGGAAGAAAAAGCTGCAGAAAAAGGAAATCGCAAGAACGTGTTGT	1800
QY	1969	GCAGAACATTTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAATTCGAATGATA	2028
Db	1801	GCAGAACATTTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACTATTGCAATGATA	1860
QY	2029	GATCGGTATCTCATCTTGGAACTTCTATATAGAAGAGAAAGATAAGAACTTTCAGTTC	2088
Db	1861	GATCGGTATCTCATCTTGGAACTTCTATATAGAAGAGAAAGATAAGAACTTTCAGTTC	1920
QY	2089	ATAGAAAGATGATAGTATGATGAGGGTGTGATGATGATGATATTGTGATGTGATGATGATGAG	2148
Db	1921	ATAGAAAGATGATAGTATGATGAGGGTGTGATGATGATGATGATATTGTGATGTGATGATGAG	1980
QY	2149	GATGATTTTAAAGAACTTTGAAACTGATGAAACAGATAGATTCTCATGACTTTATTT	2208
Db	1981	GATGATTTTAAAGAACTTTGAAACTGATGAAACAGATAGATTCTCATGACTTTATTT	2040
QY	2209	TTTGAAGAAACAATAAATGTTGAAAGAGCTGGCTGAAAAACCAGATATGAAATGAAAG	2268
Db	2041	TTTGAAGAAACAATAAATGTTGAAAGAGCTGGCTGAAAAACCAGATATGAAATGAAAG	2100
QY	2269	CTGACCAATTTAAGAAATACCATATGAGCAATATACTAGACTGAGGAATCAGACGA	2328
Db	2101	CTGACCAATTTAAGAAATACCATATGAGCAATATACTAGACTGAGGAATCAGACGA	2160
QY	2329	GGAATAATCTTTACAAAAACAGACAGAGTGCATATGCGCTTCCAGTGGATTACTGAA	2388
Db	2161	GGAATAATCTTTACAAAAACAGACAGAGTGCATATGCGCTTCCAGTGGATTACTGAA	2220
QY	2389	AATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCACTCTGATTGGAAGCTGCACACAGC	2448
Db	2221	AATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCACTCTGATTGGAAGCTGCACACAGC	2280
QY	2449	AGTGAGTTCAAAACCATGACACAGAAATGAACAAAAAGAGTCAATAGTAAATTTGCACT	2508
Db	2281	NN	2340
QY	2509	GGAAGAAATCAATCTGCTTATCGCTACACACAGTGCAGAGAAAGGTCTGGATATTAAAGAA	2568
Db	2341	GGAAGAAATCAATCTGCTTATCGCTACACACAGTGCAGAGAAAGGTCTGGATATTAAAGAA	2400
QY	2569	TGTACATTTGTTATCCGTTATGCTCTCGTCACCAATGAATAGCCATGCTCCAGGCCGT	2628
Db	2401	TGTACATTTGTTATCCGTTATGCTCTCGTCACCAATGAATAGCCATGCTNNNNNGCCGT	2460
QY	2629	GCTCGAGCCAGAGCTGATGAGAGCACTACGTCCTGCTGCTCAAGTGTTCAGAGAGTT	2688
Db	2461	GCTCGAGCCAGAGCTGATGAGAGCACTACGTCCTGCTGCTCAAGCGGTTTCAGAGAGTT	2520
QY	2689	ATCGAACATGAGACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTTGT	2748
Db	2521	ATCGAACGTCAGACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTTGT	2580
QY	2749	GTTCAAATATGAAACCAGAGAGTATGCTCATTAAGATTTTGGAAATTACAGATGCAAGT	2808
Db	2581	GTTCAAATATGAAACCAGAGAGTATGCTCATTAAGATTTTGGAAATTACAGATGCAAGT	2640
QY	2809	ATAATGGAAGAAAGAAATGAAGAACCAAGAGAAATATTCACAGCATTTACAAGAAATAAACCA	2868
Db	2641	ATAATGGAAGAAAGAAATGAAGAACCAAGAGAAATATTCACAGCATTTACAAGAAATAAACCA	2700
QY	2869	TCACTAATACTTTCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATC	2928
Db	2701	TCACTAATACTTTCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATC	2760

Oy		2929	CATGTAATTGAGAAAATGCATCACCGTCAATATGACCCAGAATTCAAGGAACCTTTACATT 	2988
Db		2761	CACGTAAATTGAGAAAAATGCATCAGTCATAATATGACCAGAAITTCAGGAACCTTTACATT 	2820
Oy		2989	GTAAGAGAAAAACAAGCACTGCCAAAAGAGTGTCGGACTATCAAAATAAATGATGAAATC 	3048
Db		2821	GTAAGAGAAAAACAAGCACTGCCAAAAGAGTGTCGGACTATCAAAATAAATGATGAAATC 	2880
Oy		3049	ATCTGCAAATGTGGCCGCGCTTGGGAAACAATGATGTCACAAAGGCTTAGATTGCGCT 	3108
Db		2881	ATCTGCAAATGTGGCCGCGCTTGGGAAACAATGATGTCACAAAGGCTTAGATTGCGCT 	2940
Oy		3109	TGCTCAAAATAAGGAATTTTGTAGTGTTTCCAATAATTCACACAAAGAACATAAC 	3168
Db		2941	TGCTCAAAATAAGGAATTTTGTAGTGTTTCCAATAATTCACACAAAGAACATAAC 	3000
Oy		3169	AAAAAGTGGGTAGAAATTAACCTATCACATTTCCCAATCTTGACTATTCAGAATGCTGTTA 	3228
Db		3001	AAAAAGTGGGTAGAAATTAACCTATCACATTTCCCAATCTTGACTATTCAGAATGCTGTTA 	3060
Oy		3229	TTTAGTGATGAGGATTAG 3246 	
Db		3061	TTTAGTGATGAGGATTAG 3078 	
RESULT 3				
AK037057				
LOCUS	AK037057	2304 bp	mRNA	linear HTC 03-APR-2004
DEFINITION	Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:9930105B04 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert sequence.			
ACCESSION	AK037057			
VERSION	AK037057.1	GI:26331913		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1			
AUTHORS	Carninci,P. and Hayashizaki,Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
PUBMED	10349636			
REFERENCE	2			
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
PUBMED	11042159			
REFERENCE	3			
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer			
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)			
PUBMED	11076861			
REFERENCE	4			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.			
TITLE	Functional annotation of a full-length mouse cDNA collection			
JOURNAL	Nature 409, 685-690 (2001)			
REFERENCE	5			
AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.			

TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 2304)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sogabe,Y., Tagami,M., Tagawa,A., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.jp/ URL:http://fantom.gsc.riken.jp/ Location/Qualifiers
FEATURES	1. 2304 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:9930105B04" /db_xref="taxon:10090" /clone="9930105B04" /sex="female" /tissue type="vagina" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 224. .>2302 /note="unnamed protein product; putative similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens] (SPTR 09H3G6, evidence: PASTY, 75.6%ID, 67.6%length, match=2079)" /codon_start=1 /protein_id="BAC29687.1" /db_xref="GI:26331914" /translation="MSIVCSAEDSFRNLILFRPRLKMYIQVEPYLDHLIFLSAETKEQILKKINTCGNTSAEELLSTLEGGQWPLGWTQMFVEALEHSGNPLARVYKPTLTDLPSPSSETAHDECLHLTLLOPTLVKLLINDVLDTCFEKGLTVEDNRISAAGNSGNESGVRELLRRIIVOKENWFSTFLDVLQGTNDALFOELTGGCGCPEDNTDLANSSHRDGPAANECLLPAYDESSLTEAMVVDILPEASCTDSVTTESPTSLAEGSVSCFDESLIGHNSNMGDSGTMGSDSDSEVIQTRVYSPPELQRLPYQMEVAQPALDGKNIICLPTGSGKTRVAYITKDHLDDKKQASGSEKVIIVLVKVMLEQLFRKEFNPYLKKMYRIIGLSGDTOLKISFPEVVKSYDVIISTAOILENSLNLNESGDDGVQSDPSLIIDECHHTNKEAVYNNIMRKYLKQKLKRNLDLKKONKPAIPLPQIIGLTASPGVGAKKQSEAEKHLINICANDAFITIKTKENLGOLKDOIKEPCKFVIADDTRENPKLEIMASIQTVCQKSPMSDFGTQHYEQWAIQMEKKAADGNRDRVCAEHLRKYNALQINDTIRMIDAYSHLETFTDEKEKKFAVLNDSKSDDEASSCNDQLKGDVKKSLKIDETDEFLMNLFFDNKKMKLKKLAEN"
ORIGIN	Query Match 41.4%; Score 1393; DB 4; Length 2304; Best Local Similarity 79.3%; Pred. No. 0; Matches 1691; Conservative 0; Mismatches 435; Indels 7; Gaps 3;
121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAAGAAAGATGCGAATGG 179	

Db 175 CCGCCCCCGCCCCGGAGACTCTCTCCCATTTCTTGAGACCTCACGATGTCGATTGT 234  
QY 180 GTATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAAAT 239  
Db 235 CTGTTCTGCAGAGGACAGCTTCAGGAATCTCATCTTAATCTTCAGGCCACGGCTGAAAT 294  
QY 240 GTACATCCAGGTGAGCCTGTGCTGAGCTAACCTGACCTTTCTGCTGACAGGTGAAGA 299  
Db 295 GTACATTCAGGTGAGCCAGTGTGACCACTCATCTTCTGTCTGCAGAAACCAAGA 354  
QY 300 GCAGATTCAGAGACAGTCCGCCACTCCGGAACTGACAGGAGTTGAATGCTGCTGAG 359  
Db 355 GCAGATTTCTTAAAAAGATCAACACTGTGTGTAACACCAAGCGGAGAACTGCTGAG 414  
QY 360 CACCTTGAGAGAGGAGTCTGCACTTGTTGGACTCGGGAATTCGTGAGGCCCTCCG 419  
Db 415 CACCTTGAGACAGGGAACAATGGCTCTTGAGATGACCGAGATGTTCTGTGAGGCCCTAGA 474  
QY 420 GAGAACCGGCAGCCCTCTGGCCGCCCTACATGAACCTGAGCTCACGGACTGCCCTC 479  
Db 475 GCACAGTGGCAATCCCTAGCCGCGCTATGTCAAACCCACACTCATGTCTGCCCTC 534  
QY 480 TGCATCGTTTGAGAACGCTCATGATGAATATCTCCAATGCTGAACTCTTCAGCCAC 539  
Db 535 TCTTCTCTGAGACTGCCCATGACGAGTGTCTCCACTGTGCTGACCCCTCTCCAGCCAC 594  
QY 540 TCTGCTGACAAGCTTCTAGTTAGACGTCCTTGAATAGTCATGAGAGGAACCTGTT 599  
Db 595 TTGCTGGACAAACTTCTGAATTAACGATGCTTGACACTTGCTTGAGAGGACTATT 654  
QY 600 GACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAATGGAATGAATCAGGTGT 659  
Db 655 GACAGTCGAAGACAGAAATCGGATTTCTGCTGCAGAAACAAGCGGAATGAGTCAGGTGT 714  
QY 660 AAGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAACCTGTTCTCTGCATTTCTGAATGT 719  
Db 715 AAGAGAGCTGCTGAGAAGGATTGTGCAGAAAGAAAACCTGTTTCTTACCTTCTGGAATGT 774  
QY 720 TCTTCTGCAAAACAGAAAACAATGAACCTGTGTCCAAGATTAAACAGCTCTGATTGCTCAGA 779  
Db 775 TCTGCGCCAAACTGGAATGATGCACTATTCCAAAGACTAAACAGGTGAGGCTGCCAGA 834  
QY 780 AAGCAATGCAGAGATTGGAATTTATCAACAATTGATGTCCTCAAGTGAAGACCAACT 839  
Db 835 AGACAAACAGACTTGCTTAATCTGCTCACAGAGATGGGCTTGCACTTAATGATGTCT 894  
QY 840 TCTTCAACACAGTTCAAGCCAAATCTGAGAAAGAGTCTGGGCGATGGAGAAATTAATC 899  
Db 895 TCTGCTGCTGTGATGAGTCAAGTCTGAGACAGAGGCTGGAACGTAGACGACATATT 954  
QY 900 ATCAGAAATCATCTTTTGCAGATTTCTTCTGTAGTTTCAGAATCAGACACAAAGTTTGGCAGA 959  
Db 955 ACCAGAGGCTTCTGTACAGATTCGTCTGTGACCAAGAAATCAGACAAAGTTTGGCAGA 1014  
QY 960 AGGAAGTGCAGCTGCTTAGATGAAGTCTTGACATTAACGCAACATGGGCGAGTATTC 1019  
Db 1015 AGGAAGTGTCACTGCTTCGATGAAGTCTTGACATTAACGCAACATGGGCGAGTATTC 1074  
QY 1020 AGGCACCATGGGAAGTGATTCAGATGAA--GAGAAATGGCAGCAAGAGCATCCCCGA 1076  
Db 1075 AGGCACCATGGGAAGTGATTCAGATGAAGTGTCAATCAACAACAAAGAGTATCCCCGA 1134  
QY 1077 GCCAGAACTCCAGCTCAGGCTTAACCAATGGAAGTTGCCAGCCAGCCTTGGAAGGAA 1136  
Db 1135 GCCAGAACTGCAGCTCAGGCTTAACCAATGGAAGTGCCCAACAGCTCTAGATGGGAA 1194  
QY 1137 GAATATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGCGTGTTHACATTC 1196  
Db 1195 GAATATTAATTAATCTGCTCCGCCAGGGGAGTGGAAAAACAGAGTGCGTGTTHACATCAC 1254  
QY 1197 CAAGATCACTTAGACAAGAAAAAAGCATCTGAGCCTGGAAGAAAGTTATAGTTCTTGT 1256  
Db 1255 CAAAGATCACTTAGACAAGAAAGCAGGCACTTGAATCCGGGAAGGTTATCGTTCTTGT 1314

QY 1257 CAATTAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACCATTTTGAAGAA 1316  
Db 1315 CAATTAAGTAAATGTAGACAACAACCTTTCCGAAAAAGATTCAACCCATATTTGAAGAA 1374  
QY 1317 ATGGTATCGTGTATTTGGATTAAAGTGTGATACCAACTGMAAATATATCATTTCCAGAAGT 1376  
Db 1375 ATGGTATCGAATTATTTGATTAAAGTGGGATACCAAGCTGMAAATATCATTTCCAGAAAGT 1434  
QY 1377 TGTCAAGTCTCTGTATATTAATATCAGTACAGCTCAATCCTTGAAAACTCCCTTAA 1436  
Db 1435 TGTCAATCTTACGATGTATTAATCAGCACTGCTCAAACTCTTGAAAACTCCCTTAA 1494  
QY 1437 CTTGAAAAATGAGAGAGATGCTGCTGTTCATATGTCAAGCTTTCCCTCATTAATCATTTGA 1496  
Db 1495 TCTGAGAGTGGAGACGATGACGGTGTGACAGCTGTCAAGCTTCTCTCATTAATCATTTGA 1554  
QY 1497 TGAATGTCATCACACCAACAAAGAGCAGTGTATTAATTAACATCATGAGGCAATATTGAT 1556  
Db 1555 TGAGTGCATCACACCAACAAAGAGGAGCAGTCTATTAACAACATCATGAGCATATTTGAA 1614  
QY 1557 GCAGAGTTGMAAAACAATAGACTCAAGMAAGMAAACAAACAGTATTCCTCTCTCA 1616  
Db 1615 GCAGAGCTGAGAAACAATGACCTCAAGMAACMAAACAGCATTCCCTGCCGA 1674  
QY 1617 GATACTGGACTAACAGCTTCACTGTGTGTGGAGGGCCACGAAGCAAGCCAAAGCTGA 1676  
Db 1675 GATACTAGGACTGACAGCTTCACTGTGTGTGGAGCCAAAGCAGTCTGAGGCTGA 1734  
QY 1677 AGAACACATTTTAAACATATGTGCCAATCTTGATGCAATTAATTAACATGTAAAGA 1736  
Db 1735 AAAACATATTTTAAATATATGTGCCAATCTTGATGCTTTAACATTAACAGTGAAGA 1794  
QY 1737 AAACCTTGATCAACTGMAAAACCAATACAGAGCCATGCCAAGAGTTGCCATTGCAGA 1796  
Db 1795 GAATCTTGCTCAACTCAACCAACCAATTAAGGAACCATGCAAGAAATTTGTGATTGCTGA 1854  
QY 1797 TGCAACGAGAGAGATCCATTTAAAGAGAACTTGAATTAATGACAAAGATTCAAAC 1856  
Db 1855 TGACACGAGAGAAATCCATTTAAAGAGAACTTGAATTAATGCAAGATTCAAGAC 1914  
QY 1857 TTATGTCAAAATGAGTCCAAATGTCAAGTTTGGAACTCAACCTATGAACAATGGGCCAT 1916  
Db 1915 TTACTGCAAAAAAGTCCAATGTCAAGTTTGGAAACCAACATTATGAGCAGTGGGCCAT 1974  
QY 1917 TCAATGGAAAAAAAGCTGCAAAAAAAGGAATGCGAAAGAACGTGTTGTGCAGAACA 1976  
Db 1975 TCAATGGAAGAAAAAAGCTGCTAAAGACGGAATCGCAAAAGATCGCGTCTGTGCAGAACA 2034  
QY 1977 TTTGAGGAAGTCAATGAGGCCCTTCAAAATTAATGACACAATTCGAATGATGATGCCGA 2036  
Db 2035 TTTGAGGAAGTCAACGAGGCCCTTCAAAATCAACGACACGATCCGAATGATGATGCATA 2094  
QY 2037 TACTCATCTGAAACTTCTTAATTAATGAAGAGAAAGTAAGAAAGTTTGACATGATGAAGA 2096  
Db 2095 TAGCCACTGGAGACATTTACACTGATGAGAAAGAAAGTTGCGAGTCT--CAA 2151  
QY 2097 TGATAGTGAAGAGGTGTGATGATGAGTATTTGTGATGATGAGATGAGATGATTT 2156  
Db 2152 TGACAGGACAAAGAGTGAAGAGGCCAGCAGTTGCAATGACCACTTAAGGCGAGTGT 2211  
QY 2157 AAAGAAACCTTTGAAACTGATGAACAAGATAGATTTCTATGACTTTATTTTGAAGAA 2216  
Db 2212 AAAGAAATCTTTGAAACTGACGAACGAGATGAATTTCTATGATTTGTTTGATTA 2271  
QY 2217 CAATTAATGTTGAAAAAGGCTGGCTGAAAAACC 2249  
Db 2272 CAAGAAAAATGTTGAAAAAACTAGCTGAAAAACC 2304

RESULT 4  
AK040519  
LOCUS AK040519 2331 bp mRNA linear HTC 03-APR-2004

DEFINITION	Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:A430105A06 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert sequence.
ACCESSION	AK040519
VERSION	AK040519.1 GI:26333792
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 Carninci, P. and Hayashizaki, Y.
AUTHORS	High-efficiency full-length cDNA cloning
TITLE	Meth. Enzymol. 303, 19-44 (1999)
JOURNAL	10349636
PUBMED	
REFERENCE	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
AUTHORS	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
TITLE	Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL	11042159
PUBMED	
REFERENCE	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
PUBMED	
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
PUBMED	6 (bases 1 to 2331)
REFERENCE	6
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:ftp://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL:ftp://genome.gsc.riken.jp/  
URL:ftp://fantom.gsc.riken.jp/.

FEATURES

source

1..2331  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:A430105A06"  
/db\_xref="taxon:10090"  
/clone="A430105A06"  
/issue\_type="thymus"  
/clone\_1lb="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="0 day neonate"  
283..>2331  
/note="unnamed protein product; putative  
similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5  
[Homo sapiens] (SPTK|Q9H3G6, evidence: PASTY, 75.6%ID,  
67.6%length, match=2079)"  
/codon\_start=1  
/protein\_id="BAC30614.1"  
/db\_xref="GI:26333793"  
/translation="MSIVCSAEDSFENLILFFRPRLKMYIOVEPVLDDLFLSAETKE  
QILKINTCGNTSAAELLSTLEOGWPLGWTQMFEALHNSGNPLARYVKPTLTDL  
PSPSETHDECLHLTLLOPTVDKLLINDVLDTCPEKGLTVEDRNRIISAAGNSG  
ESGRELRLRIVOKENWFSTFLDVRQTNDALFQELTGGGCPBDNTDLANSHRDGP  
AANECPLPAYDESSLETAENVVDILPEASCTDSVTTESDPSLAESVSCFDESIGH  
NSNMGRLDSGTWGSDESIVIQTKRVSPEPELQLRPYQMEVAQPALDGKNIICLPTGS  
GKTRVAVYITRDHLDDKKQASESGRYIVLVNKMVLAEOFLREKFNYPYLKKMYRIIGLS  
GDTQLKISPEVWKSVDYIISTAOILNLSLNLIESGDDGVQVLSLIIDECHHTN  
KEAVNNIMRRLYLKQRLRNNDLKKQNKPAIPLPQILGLTASGVGAACKQSEAEKHIL  
NICANLDAFTIKTVKENLGLKHQIKPECKKFFVIADDTRENPFEKLLLEIMASIQTVC  
QKSPMSDFGTQHYEQWAIQMEKKAADGNRKDRVCAEHLRKRYNEALQINDTIRMIDAY  
SHLETFTYDEKEKKFAVLNDSDKSDDEASSCNDQLKGDVYKSLKLDETDEFLMNLFPD  
N"

CDS

ORIGIN  
Query Match 40.6%; Score 1367.4; DB 4; Length 2331;  
Best local Similarity 79.2%; Pred. No. 0;  
Matches 1663; Conservative 0; Mismatches 431; Indels 7; Gaps 3;  
QY 121 CAGCAACCATCTGCTTGGAGAAACCTCTCCC-TTCTTGAGAAAGAAGATGCGAATGG 179  
DB 234 CCGCCCCCGCCGCCGGAGACTCTCTCCATTTCCTGAGAACCTCAGCATGTCGATTGT 293  
QY 180 GTATTCACAGACGAGAAATTCGCGTATCTCATCTCGTCTTCAAGGCCAGGGTGAAAT 239  
DB 294 CTGTTCTGACAGAGACAGCTTCAGGAATCTATCTTATTCTTCAGGCCAGGCTGAAAT 353  
QY 240 GTACATCCAGGTGAGCGCTGTGCTGACTACCTGACCTTTCTGCGCTGACAGAGGTGAAGA 299  
DB 354 GTACATTCAGGTGAGCGAGTGTGACACCACTCATCTTTCTGTGTGAGAAACCAAGA 413  
QY 300 GCAGATTCAGAGACAGTCGCCACCTCCGGGAACATGACGACAGTGAAGTGTGCTGAG 359  
DB 414 GCAGATTTTAAAAAGATCAACACCTGTGTGTAACACCAAGCGGACAGAACTGTGCTGAG 473  
QY 360 CACCTTGAGAGAGGAGTCTGGCACCTTGCTTGAACCTGGGAATTCGTGAGAGCCCTCCG 419  
DB 474 CACCTTGAGAGAGGAGCAATGGCCTCTGGAGATGACGCAAGATGTTCTGTGAGAGCCCTAGA 533  
QY 420 GAGAACCGGACGCCCTCTGGCGCCGCCGCTACATGAACCTGAGCTACGGACTTGCCCTC 479  
DB 534 GCACAGTGGCAATCCCTTAAGCCGCGCTATGTCAAAACCACTACATGATCTGCCCTC 593  
QY 480 TCCATCGTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACCTCTTCAGCCAC 539  
DB 594 TCCTTCCTCTGAGACTGCCCATGACGAGTGTCTCACTTGCTGACCCCTCTCCAGCCAC 653  
QY 540 TCTGTGAGACAAGCTTCTAGTTAGAGAGCTTGGATAAGTGATGAGAGGAAGAACTGTT 599  
DB 654 TTGTGTGAGACAACCTTCTAGTTAAGATGTCTTGACACTTGCTTGAAGAGGAAGAACTATT 713

QY	600	GACAATTGAAGACAGAAACCGGATTGCTGTCGAGAAAACAATGMAATGAAATCAGGTGT	659
Db	714	GACAGTCGAAGACAGAAATCGGATTTCTGCTGCAGGAAACAGCGGAATGAGTCAGGTGT	773
QY	660	AAGAGAGCTACTAATAAAGGATTGTGCAGAAAGAAAATGCTGTTCTCTGCATTTCTGAATGT	719
Db	774	AAGAGAGCTGTGAGAGGATTTGTGAGAAAGAAAATGCTTTCTACTCTTCTGAGATGT	833
QY	720	TCTTCGTAACACAGAAACAATGAATCTTGTCCAAGATTAAACAGGCTTGATTTGCTCAGA	779
Db	834	TCTGCGCCAACTGGAATGATGCACTATTTCCAGAACTAACAGGTGAGGCTGCCCCAGA	893
QY	780	AAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGTCTCTCAAGTGAAGACCACT	839
Db	894	AGACAAACACAGACTTGGCTTAATCTGCTCTCACAGAGATGGGCTTCAGCTTAATGATGTCT	953
QY	840	TCTTTCAAACCAACAGTTCAGCCAAATCTGAGAGAGAGGTCTGGGGCATGAGAAATACTC	899
Db	954	TCTGCTGTGTGATGAGTCAAGTCTGAGACAGAGGCTTGAACGTAGACGACATATT	1013
QY	900	ATCAGAATCATCTTTTCGAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGA	959
Db	1014	ACCAGAGGCTTCTGTACAGATTCTGTGTGACCAAGAAATCAGACAAAGTTTGGCAGA	1073
QY	960	AGGAAGTGCAGCTGCTTAGATGAAGTCTTGGACATTAACAGCAACATGGGCGAGTATTC	1019
Db	1074	AGGAAGTGCAGCTGCTTCGATGAAGTCTTGGACATAACAGCAACATGGGCGAGGATTC	1133
QY	1020	AGGCACCATGGGAAGTGAATTCAGATGAA--GAGAACTGGCAGCAAGAGCATCCCCGA	1076
Db	1134	AGGCACCATGGGAAGTGAATTCAGATGAAGTGTCTATCCAGACAAAAAGATATCCCCGA	1193
QY	1077	GCCAGAACTCCAGCTCAGGCTTACCMAATGGAAGTTGCCCAAGCCAGCTTGGAAAGGAA	1136
Db	1194	GCCAGAACTGCAGCTCAGGCTTACCMAATGGAAGTGGCCCAACCAAGCTCTAGATGGGAA	1253
QY	1137	GAATATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGCGTGTTTACATTGC	1196
Db	1254	GAATATTATTATCTGCTCCCAAGGAGTGGAAAAACAGAGTGCGTGTTTACATCAC	1313
QY	1197	CAAGATCACTTAGACAAAGAAAAAGCATCTGAGCCTGGAAAAATTATAGTCTTGT	1256
Db	1314	CAAAAGTCACTTAGACAAAGAAAGCAGGCATCTGAATCCGGAAGTTATCGTCTTGT	1373
QY	1257	CAATAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACCATTTTGAAGAA	1316
Db	1374	CAATAAGTAAATGTTAGCAGAACAACTTTCCGAAAGAGTTCAACCCATATTTGAAGAA	1433
QY	1317	ATGCTATCGTGTATTGGATTAAAGTGCGTATACCCAATGAATAATCATTTCCAGAAAT	1376
Db	1434	ATGCTATCGAATTATTGGATTAAAGTGCGGATACCCAGCTGAATAATATCATTTCCAGAAAT	1493
QY	1377	TGTCAAAGTCCGTGATATTATTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTTTAAA	1436
Db	1494	TGTCAAATCTTAGATGTTATTATCAGCACTGCTCAAAATCCTTGAAAACTCCCTTTAAA	1553
QY	1437	CTTGAAAAATGAGAAAGATGCTGTGTTCAATTGTCAAGCTTTTCCCTCATTTATGATGA	1496
Db	1554	TCTGGAAGTGGAGACGATGACGGTGTGCAAGCTGTCAAGCTTCTCTCATTTATGATGA	1613
QY	1497	TGAATGTCATCACACCAACAAAGAACAGTGTATATAATCATCATGAGGCATTATTGAT	1556
Db	1614	TGAGTGCCATCACACCAACAGAGAGGCACTATTAACAACATCATGAGACGATATTGAA	1673
QY	1557	GCAGAAAGTTGAAAAACAATAGACTCAAGAAAAAGAAAAACAACCAAGTATTCCTCTCTCA	1616
Db	1674	GCAGAAAGCTGAGAAAAACAATGACTCAAGAAAAACAACCAAGCCATTCCCTGCGCA	1733
QY	1617	GATACTGGGACTAACAGCTTCACTGTGTTGGAGGGGCCAGAAAGCAAGCCAAAGCTGA	1676
Db	1734	GATACTAGGACTGACAGCTTCACTGTGTTGGAGCAGCCAAAAAGCAGTCTGAGGCTGA	1793
QY	1677	AGAAACATTTTAAACTATGTGCCAATCTTGATGCAATTAATAAACTGTTAAAGA	1736

Db	1794	AAAACATATTTTAAATATATATATGTGCCAATCTTGATGCCCTTTACCATTTAAACAGTGAAGA	1853
Qy	1737	AAACCTTGATCAACTGAAAAAACCAATACAGAGCCATGCAAGAGTTGCCATTGCAGA	1796
Db	1854	GAATCTTGTCACACTCAAAACACCAATATAAGAAACCATGCAAGAAATTTGTGATTGCTGA	1913
Qy	1797	TGCAACCAGAGAGATCCATTTAAAGAGAACTTCTAGAAATATGACAAAGATTCAAA	1856
Db	1914	TGACACCAGAGAAATCCATTTAAAGAGAACTTCTAGAAATATGCAAGCATTCAGAC	1973
Qy	1857	TTATTGTCAAAATGAGTCCCAATGTCAGATTTTGGAACTCAACCCATAGAACATGGGCCAT	1916
Db	1974	TTACTGCCAAAAAAGTCCCAATGTCAGATTTTGGAAACCAACATTATGAGCAGTGGCCAT	2033
Qy	1917	TCAAATGAAAAAAAAGCTGCAAAAAAAGAGAAATCGCAAGAAGCCTGTTGTGCAGAAC	1976
Db	2034	TCAAATGAGAAAAAAGCTGCTAAAGACGGAAATCGCAAGAATCGCGTCTGTGCAGAAC	2093
Qy	1977	TTTGAGGAAGTACAAATGAGGCCCTCAAAATTAATGACACAATTCGAATGATGATGCGTA	2036
Db	2094	TTTGAGGAAGTACAAACGAAGCCCTCAAAATCAACGACAGATCCGAATGATGATGCATA	2153
Qy	2037	TACTCATCTTGAACTTCTATATGAAGAGAAAGATAAGAGTTTGCACTCATAGAAGA	2096
Db	2154	TAGCCACCTGAGACATTTCTACACTGATGAGAAAGAAAAGAGTTCGCAATCCT---CAA	2210
Qy	2097	TGATAGTATGATGAGGCTGTGATGATGAGTATTTGTATGATGATGATGAGATGATTT	2156
Db	2211	TGACAGCGACAAAGTATGATGACGAGGCCACGATTCGAATGACCAACTTAAGGGCGATGT	2270
Qy	2157	AAAGAAACCTTTGAAACTGCGATGAACACAGATGATTTCTCATGACTTTATTTTGA	2216
Db	2271	AAAGAAATCTTTGAAACTGCGACGAAACGATGAATTTCTCATGATTTGTCTTGTATTA	2330
Qy	2217	C 2217	
Db	2331	C 2331	

```

RESULT 5
BQ233683
LOCUS
DEFINITION BQ233683 1013 bp mRNA linear EST 02-MAY-2002
AGENCOCURT_7258428 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5786526
5', mRNA sequence.
ACCESSION BQ233683
VERSION BQ233683.1 GI:20415083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1013)
AUTHORS NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM12875 row: g column: 07
High quality sequence stop: 721.
Location/Qualifiers
1..1013
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

/clone="IMAGE:5786526"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb. "

ORIGIN

Query Match		27.5%;	Score 926.4;	DB 3;	Length 1013;
Best Local Similarity		98.3%;	Pred. No. 8.4e-215;		
Matches 989;		Conservative 0;	Mismatches 11;	Indels 6;	Gaps 5;
QY	1752	GAAAAACCAATACAGAGCCATGCAAGAAGTTGCCATTCAGATGCAACCAAGAGAAGA 1811			
DB	1	GAAAAACCAATACAGAGCCATGCAAGAAGTTGCCATTCAGATGCAACCAAGAGAAGA 60			
QY	1812	TCCATTTAAAGAGAACTTCTAGAAATAATGACAAGAAATCAAACTTATTGTCAATGAG 1871			
DB	61	TCCATTTAAAGAGAACTTCTAGAAATAATGACAAGAAATCAAACTTATTGTCAATGAG 120			
QY	1872	TCCAAATGTCAGATTTTGGAACTCAACCCATATGAACAATGGGCCATTCAAAATGAAAAAAA 1931			
DB	121	TCCAAATGTCAGATTTTGGAACTCAACCCATATGAACAATGGGCCATTCAAAATGAAAAAAA 180			
QY	1932	AGCTGCAAAAAAGAAATCGCAAGAAGCGTGTGTCGAGAACATTTGAGGAAGTACAA 1991			
DB	181	AGCTGCAAAAAAGAAATCGCAAGAAGCGTGTGTCGAGAACATTTGAGGAAGTACAA 240			
QY	1992	TGAGGCCCTACAAATTAATGACACAATTGCAATGATAGATGCGTATACTCATCTTGAAC 2051			
DB	241	TGAGGCCCTACAAATTAATGACACAATTGCAATGATAGATGCGTATACTCATCTTGAAC 300			
QY	2052	TTTCTAATGAAGAGAAAGATTAAGAGTTTGCACTCATAGAAGATGATAGTGAAGG 2111			
DB	301	TTTCTAATGAAGAGAAAGATTAAGAGTTTGCACTCATAGAAGATGATAGTGAAGG 360			
QY	2112	TGGTGATGATGAGTATTGTGATGCGTGAATGAAGATGAGGATTTAAAGAAACCTTGAA 2171			
DB	361	TGGTGATGATGAGTATTGTGATGCGTGAATGAAGATGAGGATTTAAAGAAACCTTGAA 420			
QY	2172	ACTGATGAAACAGATAGATTCTCATGACTTTATTTTTGA AAAACAATAAAATGTTGAA 2231			
DB	421	ACTGATGAAACAGATAGATTCTCATGACTTTATTTTTGA AAAACAATAAAATGTTGAA 480			
QY	2232	AAGGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAATTAAGAAATACCAT 2291			
DB	481	AAGGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAATTAAGAAATACCAT 540			
QY	2292	AATGAGCAATATACTAGAGCTGAGGAATCAGCACGAGGAATTAATCTTTACAAAAACACG 2351			
DB	541	AATGAGCAATATACTAGAGCTGAGGAATCAGCACGAGGAATTAATCTTTACAAAAACACG 600			
QY	2352	ACAGAGTCATATGCGCTTTCCAGTGAATTAAGAAAAATTTGCTGAAGTAGG 2411			
DB	601	ACAGAGTCATATGCGCTTTCCAGTGAATTAAGAAAAATTTGCTGAAGTAGG 660			
QY	2412	AGTCAAAAGCCCAACCATCTGATTGGAGCTGACACAGCAGTGAAGTTCAAAACCATGACACA 2471			
DB	661	AGTCAAAAGCCCAACCATCTGATTGGAGCTGACACAGCAGTGAAGTTCAAAACCATGACACA 720			
QY	2472	GAATGAACAAAAAGAGTCATTAAGTAATTTGCGACATGGA AAAATCAATCTGCTTATCGC 2531			
DB	721	GAATGAACAAAAAGAGTCATTAAGTAATTTGCGACATGGA AAAATCAATCTGCTTATCGC 780			
QY	2532	TACCAAGTGGCAGAGAAGGCTGGATATTTAAAGAAATGAACATTTGTTATCCGTTATGG 2591			
DB	781	TACCAAGTGGCAGAGAAGGCTGGATATTTAAAGAAATGAACATTTGTTATCCGTTATGG 840			
QY	2592	TCTGCTCAACCAATGAATAGCATTGCTCCAGGCCCGTGTGAGCCAGAGCTGATGAGAG 2651			
DB	841	TCTGCTCAACCAATGAATAGCATTGCTCCAGGCCCGTGTGAGCCAGAGCTGATGAGAG 899			

QY	2652	CACCTACGCTCTGTTGCTCAGAGTGGTTCAGAGATTATGCAACATGAGACAGTTAATGA 2711			
DB	900	CACCTACGCTCTGTTGCTCAGAGTGGTTCAGAGATTATGCAACATGAGACAGTTAATGA 956			
QY	2712	TTTCCGAGAGAAGATGATGTATTAAGCTATACATTTGTTCAAAAT 2757			
DB	957	TTTCCGAGA-AAGATGATGTAT-AAGCTATACATGGGGTTCAAAAT 1000			

RESULT 6  
BU902097 1174 bp mRNA linear EST 17-OCT-2002  
LOCUS AGENCOURT\_10127740 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6502757  
DEFINITION 5', mRNA sequence.  
ACCESSION BU902097  
VERSION BU902097.1 GI:24084010  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.

REFERENCE 1 (bases 1 to 1174)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LIML at:  
http://image.llnl.gov  
Plate: LIML4059 row: f column: 06  
High quality sequence stop: 808.

FEATURES

source location/Qualifiers  
1..1174  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6502757"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb. "

ORIGIN

Query Match		26.8%;	Score 900.4;	DB 5;	Length 1174;
Best Local Similarity		93.7%;	Pred. No. 2e-208;		
Matches 1049;		Conservative 0;	Mismatches 56;	Indels 15;	Gaps 10;
QY	1	GCGCGCGGCTGAGAGCCCTGTGACAACCTCGTCAATTGTCAGGCACAGCGGTAGAC 60			
DB	55	GCGCGCGGCTGAGAGCCCTGTGACAACCTCGTCAATTGTCAGGCACAGCGGTAGAC 114			
QY	61	CCTGCTCTTAAGTGGCAGCGGACAGCGGACGACATTTCACCTGTCCGACAGACAA 120			
DB	115	CCTGCTCTTAAGTGGCAGCGGACAGCGGACGACATTTCACCTGTCCGACAGACAA 174			
QY	121	CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGCAATGGG 180			
DB	175	CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGCAATGGG 234			
QY	181	TATTCCACAGACGAGAATTTCCGCTATCTCATCTGCTTTCAGGGCCAGGGTGAATG 240			
DB	235	TATTCCACAGACGAGAATTTCCGCTATCTCATCTGCTTTCAGGGCCAGGGTGAATG 294			
QY	241	TACATCCAGGTGAGCGCTGTGCTGACTACCTGACCTTTCTGCTGCAAGGTGAAGAG 300			

Db 295 TACATCCAGGTGAGCCTTGCTGGACTACCTGACCTTTCTGCTGCAGAGGTGAAGAG 354  
Qy 301 CAGATTCAGAGGACAGTCCGCACTCCGGGAACATGCAGGCACTGAACTGCTGAGC 360  
Db 355 CAGATTCAAGAGACAGTCCGCACTCCGGGAACATGCAGGCACTGAACTGCTGAGC 414  
Qy 361 ACCTTGAGAGAGGAGTCTGGCACTTGGTGGACTCGGAAATTCGTGAGGCCCTCCGG 420  
Db 415 ACCTTGAGAGAGGAGTCTGGCACTTGGTGGACTCGGAAATTCGTGAGGCCCTCCGG 474  
Qy 421 AGAACCGGAGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCAGGACTTGCCCTCT 480  
Db 475 AGAACCGGAGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCAGGACTTGCCCTCT 534  
Qy 481 CCATCGTTTGAGAACGCTCATGATGATATCTCCAAGTGTGAACCTCCTTCAGCCCACT 540  
Db 535 CCATCGTTTGAGAACGCTCATGATGATATCTCCAAGTGTGAACCTCCTTCAGCCCACT 594  
Qy 541 CTGGTGACAAGCTTCTAGTGAAGACGCTCTTGGAATAGTCATGAGAGGAACTGTTG 600  
Db 595 CTGGTGACAAGCTTCTAGTGAAGACGCTCTTGGAATAGTCATGAGAGGAACTGTTG 654  
Qy 601 ACAATTGAAGACAGAAACCGGATGCTGCTGCAGAAAACAATGAAATGAATCAGGTGA 660  
Db 655 ACAATTGAAGACAGAAACCGGATGCTGCTGCAGAAAACAATGAAATGAATCAGGTGA 714  
Qy 661 AGAGAGCTACTAAAAAGATTGTGCAGAAAAAACTGTTCTCTGCATTTCTGAATGTT 720  
Db 715 AGAGAGCTACTAAAAAGATTGTGCAGAAAAAACTGTTCTCTGCATTTCTGAATGTT 774  
Qy 721 CTTCGTCAACACAGAAACAATGAACTTGTCCAAGATTAAACAGGCTCTGATTGCTCAGAA 780  
Db 775 CTTCGTCAACACAGAAACAATGAACTTGTCCAAGATTAAACAGGCTCTGATTGCTCAGAA 834  
Qy 781 AGCAATGCAGAGATTGAGAAATTATCAACAAGTTGATGTTCTCAAGT-GGAAAGCAACT 839  
Db 835 AGCAATGCAGAGATTGAGAAATTATCAACAAGTTGATGTTCTCAAGTGGGAAAAAGCAACT 894  
Qy 840 TCTTTCAACACACA-GTTCAGCCAAATCTGG-AGAAGAGGCTCGGGGCATGG-AGAATTA 896  
Db 895 TCTTTCAACACACAGTTTCAAGCCAAATCTGAAAAAGAGTCTGGGCATGAAAAATTAA 954  
Qy 897 CTCATC-AGAATCATCTTTTGAGATTTCTTCTGTAGTTTCAGAAATCAGACACAAGTTGG 955  
Db 955 CTCATCAAAAATCATCTTTTGAGATTTCTTCTGTAGTTTCAGAAACCAAAACAAGTTGGG 1014  
Qy 956 C-AGAAGAAAGTGTAGCTGCTT-AGATGAAAGTCTT-GGACATAACAGCAATGGGCA 1012  
Db 1015 CAAAAAGAAATGGCACTGCTTAATAATGAAGTCTTGGGACATACAAACAATGGGCA 1074  
Qy 1013 GTGATTCAGGCACCATGGG---AAGTATTCAGATGAAGAATGTGCGACAGAGACA- 1068  
Db 1075 GGGATTTCCGGCCCCCGGGGCAAGATTTTCAATTAATAAAAAAAGGGGCCACCAAGAAA 1134  
Qy 1069 ---TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCATA 1105  
Db 1135 ATTCCCGGGAACAAAAATTCCAGCTTGGGCCCTTAACAAA 1174

RESULT 7  
BM467983 1115 bp mRNA linear EST 05-FEB-2002  
LOCUS BM467983  
DEFINITION AGENCOURT 6437921 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5532884  
5', mRNA sequence.  
ACCESSION BM467983  
VERSION BM467983.1 GI:18517025  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

REFERENCE 1 (bases 1 to 1115)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12216 row: n column: 21  
High quality sequence stop: 690.  
location/Qualifiers

FEATURES  
source  
1. 1115  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5532884"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: PCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb."

ORIGIN

Query Match 24.0%; Score 806.8; DB 3; Length 1115;  
Best Local Similarity 99.6%; Pred. No. 1.5e-185;  
Matches 808; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2555 TGGATATTAAAGATGTAAACATTTGTTATCCGTTATGCTCTCGTCACCAATGAATAGCCA 2614  
Db 2 TGGATATTAAAGATGTAAACATTTGTTATCCGTTATGCTCTCGTCACCAATGAATAGCCA 61  
Qy 2615 TGGTCCAGGCCCGTGTGCGAGCCAGAGCTGATGAGACACCTACGTCCTGGTTGCTCACA 2674  
Db 62 TGGTCCAGGCCCGTGTGCGAGCCAGAGCTGATGAGACACCTACGTCCTGGTTGCTCACA 121  
Qy 2675 GTGGTTCAAGAGTTATCGAACATGAGACAGTTAATGATTTCCGAGAGAAAGATGTATTA 2734  
Db 122 GTGGTTCAAGAGTTATCGAACGTGAGACAGTTAATGATTTCCGAGAGAAAGATGTATTA 181  
Qy 2735 AAGCTATACATTGTGTTCAAAATATGAACCAGAGAGTATGCTCATTAAGATTTTGGAAAT 2794  
Db 182 AAGCTATACATTGTGTTCAAAATATGAACCAGAGAGTATGCTCATTAAGATTTTGGAAAT 241  
Qy 2795 TACAGATGCAAAAGTATATGAGAAAAAGAAATGAACCAAGAAATATGGCAAGCATT 2854  
Db 242 TACAGATGCAAAAGTATATGAGAAAAAGAAATGAACCAAGAAATATGGCAAGCATT 301  
Qy 2855 ACAAGATAACCCATCACTAATACTTCTCTTGGCAAAAAGTGCAGTGTGCTAGCCTGTT 2914  
Db 302 ACAAGATAACCCATCACTAATACTTCTCTTGGCAAAAAGTGCAGTGTGCTAGCCTGTT 361  
Qy 2915 CTGGGGAAGATATCCATGTATAATTGAGAAATGCATCAGTCAATATGACCCGAGAATTCA 2974  
Db 362 CTGGGGAAGATATCCATGTATAATTGAGAAATGCATCAGTCAATATGACCCGAGAATTCA 421  
Qy 2975 AGGAATTTTACATTGTGAAGAAAAAACAAGCACTGCAAAAAGAAAGTGGCCGACTATCAAA 3034  
Db 422 AGGAATTTTACATTGTGAAGAAAAAACAAGCACTGCAAAAAGAAAGTGGCCGACTATCAAA 481  
Qy 3035 TAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAAG 3094  
Db 482 TAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAAG 541  
Qy 3095 GCTTAGATTGGCTTGTCTCAAAATAAGGAATTTTGTAGTGTTTCAAAAATAATTCAA 3154  
Db 542 GCTTAGATTGGCTTGTCTCAAAATAAGGAATTTTGTAGTGTTTCAAAAATAATTCAA 601  
Qy 3155 CAAAGAAACAATACAAAAAGTGGGTGAATTACCTATCACATTTCCCAATCTTGACTATT 3214

|||||  
Db 602 CAAGAACAATACAAAAAGGGGTAGATTACCTATCACATTTCCCAATCTGACTATT 661  
Qy 3215 CAGATGCTGTTTATTAGTGAAGGATTAGCAGCTTGATTTGAAGATTCTTTAAATAC 3274  
Db 662 CAGAACTGCTTTATTATTAGTGAAGGATTAGCAGCTTGATTTGAAGATTCTTTAAATAC 721  
Qy 3275 TATCAGTTAAACATTAAATATGATTATGATTAAATGATTATCTATGCTACAGAACTGACA 3334  
Db 722 TATCAGTTAAACATTAAATATGATTATGATTAAATGATTATCTATGCTACAGAACTGACA 781  
Qy 3335 TAAGATCAATAAAATGATTGTTTACTCTG 3365  
Db 782 TAAGATCAATAAAATGATTGTTTACTCTG 812  
RESULT 8  
BQ960157 870 bp mRNA linear EST 21-AUG-2002  
LOCUS AGENCOURT\_8949486 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6472264  
DEFINITION 5', mRNA sequence.  
ACCESSION BQ960157  
VERSION BQ960157.1 GI:22375635  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 870)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM14005 row: 0 column: 17  
High quality sequence stop: 738.  
FEATURES  
source location/Qualifiers  
1..870  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6472264"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1ib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb."  
ORIGIN  
Query Match 23.8%; Score 801.8; DB 5; Length 870;  
Best Local Similarity 98.4%; Pred. No. 2.3e-184;  
Matches 851; Conservative 0; Mismatches 9; Indels 5; Gaps 4;  
Qy 1940 AAAAAAGAAATCGCAAGAAGCTGTTGTGCAAGACATTTGAGGAAGTACAAATGAGGCC 1999  
Db 1 AAGAAGAAATCGCAAGAAGCTGTTGTGCAAGAA-ATTGAGGAAGTACAAATGAGGCC 59  
Qy 2000 TACAAATTAAATGACACAATTCGAATGATGATGCGTATACTCATCTTGAACCTTTCTATA 2059  
Db 60 TACAAATTAAATGACACAATTCGAATGATGATGCGTATACTCATCTTGAACCTTTCTATA 119  
Qy 2060 ATGAAGAGAAAGATAGAAGTTTGCAGTCATAGAAGATGATAGTGAGGGGTGTGATG 2119  
Db 120 ATGAAGAGAAAGATAGAAGTTTGCAGTCATAGAAGATGATAGTGAGGGGTGTGATG 179

Qy 2120 ATGAGTATTGTGATGGTGAAGATGAGATGATTTAAAGAAACCTTGAACCTGATG 2179  
Db 180 ATGAGTATTGTGATGGTGAAGATGAGATGATTTAAAGAAACCTTGAACCTGATG 239  
Qy 2180 AAACAGATAGATTTCATGACTTTATTTTGA AAAACAATAAATGTTGA AAAAGCTGG 2239  
Db 240 AAACAGATAGATTTCATGACTTTATTTTGA AAAACAATAAATGTTGA AAAAGCTGG 299  
Qy 2240 CTGAAAAACCCAGAAATGAAAAATGAAAAGCTGACCAATTAAGAATACCAATATGAGC 2299  
Db 300 CTGAAAAACCCAGAAATGAAAAATGAAAAGCTGACCAATTAAGAATACCAATATGAGC 359  
Qy 2300 AATATCTAGAGCTGAGGAATCAGCAGGAAATTAATCTTTACAAAAACGACAGAGTG 2359  
Db 360 AATATCTAGAGCTGAGGAATCAGCAGGAAATTAATCTTTACAAAAACGACAGAGTG 419  
Qy 2360 CATATGCGCTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGCAAG 2419  
Db 420 CATATGCGCTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGCAAG 479  
Qy 2420 CCCACCATCTGATTGGAGCTGACACAGCAGTGAATTTCAAACCCATGACACAGATGAC 2479  
Db 480 CCCACCATCTGATTGGAGCTGACACAGCAGTGAATTTCAAACCCATGACACAGATGAC 539  
Qy 2480 AAAAAAGATCATTAGTAAATTTGCGACTGAAAAATCAATCTGCTTATGCTACCAAG 2539  
Db 540 AAAAAAGATCATTAGTAAATTTGCGACTGAAAAATCAATCTGCTTATGCTACCAAG 599  
Qy 2540 TGGCAGAAAGAGTCTGGATTAATAAGAAATGTAACATTTGTTATCCGTTATGCTCTGCA 2599  
Db 600 TGGCAGAAAGAGTCTGGATTAATAAGAAATGTAACATTTGTTATCCGTTATGCTCTGCA 659  
Qy 2600 CCAATGAATAGCCATGGTCCAGGCCCGTGGTGCAGCCAGAGCTGATGAGCAGCACTAGC 2659  
Db 660 CCAATGAATAGCCATGGTCCAGGCCCGTGGTGCAGCCAGAGCTGATGAGCAGCACTAGC 719  
Qy 2660 TCCTGTTGCTCAGAGT-GGTTTCAGAGATTATGCAACATGAGACAGTTATGATTTCCGA 2718  
Db 720 TCCTGTTGCTCAGAGTGGTTTCAGAGATTATGCAACGTGAGACAGTTATGATTTCCGA 779  
Qy 2719 GAGAAATGATGTATTAAGCTATACATTTGTTCAAAATATG-AAACCAAGAGATAT-- 2775  
Db 780 GAGAAATGATGTATTAAGCTATACATTTGTTCAAAATATGAAAAACCAAGAGATATG 839  
Qy 2776 GCTCATTAAGATTTTGAATTAACA 2800  
Db 840 GCTCATTAAGATTTTGAATTAACA 864  
RESULT 9  
BUI89982 918 bp mRNA linear EST 04-SEP-2002  
LOCUS AGENCOURT\_7950921 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6173848  
DEFINITION 5', mRNA sequence.  
ACCESSION BUI89982  
VERSION BUI89982.1 GI:22703966  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 918)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation



QY 2690 TCGAATGAGACAGTAAATGATTTCCGAGAGAAGATGATGTATAAGCTATACATTGTG 2749  
| | | | |  
Db 695 TCGAACATGAGACAGTAAATGATTTCCGAGAGAAGATGATGTATAAGCTATACATTGTG 636  
QY 2750 TTCAAAATATGAAACCAGAGGATATGCTCATAGATTTTGGAAATTACAGTGCAGAGTA 2809  
| | | | |  
Db 635 TTCAAAATATGAAACCAGAGGATATGCTCATAGATTTTGGAAATTACAGTGCAGAGTA 576  
QY 2810 TAAAGAAAAAGAAATGAAAAACAAGAAATATTTGCCAAGCATTACAGAAATAACCCAT 2869  
| | | | |  
Db 575 TAAAGAAAAAGAAATGAAAAACAAGAAATATTTGCCAAGCATTACAGAAATAACCCAT 516  
QY 2870 CACTAATACTTTCTCTTGGAAAACTGCAGTGTGCTAGCCTGTTCTGGGAGAGATATCC 2929  
| | | | |  
Db 515 CACTAATACTTTCTCTTGGAAAACTGCAGTGTGCTAGCCTGTTCTGGGAGAGATATCC 456  
QY 2930 ATGTAATTGAGAAAAATGCATCAGTCAATATGACCCCAGAATTCAAGGAATTACATTG 2989  
| | | | |  
Db 455 ATGTAATTGAGAAAAATGCATCAGTCAATATGACCCCAGAATTCAAGGAATTACATTG 396  
QY 2990 TAAAGAAAAACAAGCACTGCAAAAAGAAGTGTGCCGACTATCAATTAATGTGAATCA 3049  
| | | | |  
Db 395 TAAAGAAAAACAAGCACTGCAAAAAGAAGTGTGCCGACTATCAATTAATGTGAATCA 336  
QY 3050 TCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAAGCTTAGATTGCCCTT 3109  
| | | | |  
Db 335 TCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAAGCTTAGATTGCCCTT 276  
QY 3110 GTCTCAAAATAAGGAATTTTGTAGTGTCTTCAAAAATAATTCACAAAGAAACAATACA 3169  
| | | | |  
Db 275 GTCTCAAAATAAGGAATTTTGTAGTGTCTTCAAAAATAATTCACAAAGAAACAATACA 216  
QY 3170 AAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTTTAT 3229  
| | | | |  
Db 215 AAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTTTAT 156  
QY 3230 TTAATGATGAGGATTAGACATTGATGTAATCATTTTAAATACTATCAGTTAAACATT 3289  
| | | | |  
Db 155 TTAATGATGAGGATTAGACATTGATGTAATCATTTTAAATACTATCAGTTAAACATT 96  
QY 3290 TAAATGATGATGATTAATGATTAATCATTTATGCTACAGAACTGACATAAGAAATCAATAAAA 3349  
| | | | |  
Db 95 TAAATGATGATGATTAATGATTAATCATTTATGCTACAGAACTGACATAAGAAATCAATAAAA 36  
QY 3350 TGATGTTTTTACTCTG 3365  
| | | | |  
Db 35 TGATGTTTTTACTCTG 20

RESULT 11  
BF337464 1035 bp mRNA linear EST 22-NOV-2000  
LOCUS  
DEFINITION 602035195F1 NCI\_CGAP\_Brn64 Homo sapiens cDNA clone IMAGE:4183126  
5', mRNA sequence.  
ACCESSION BF337464  
VERSION BF337464.1 GI:11283715  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
1 (bases 1 to 1035)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
plate: LLM9498 row: n column: 23  
High quality sequence start: 5  
High quality sequence stop: 695.  
location/Qualifiers  
1. 1035  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4183126"  
/tissue\_type="glioblastoma with EGFR amplification"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Brn64"  
/note="Organ: brain; Vector: pCMV-Sport6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.57 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP library."  
ORIGIN  
Query Match 21.1%; Score 711.6; DB 2; Length 1035;  
Best Local Similarity 97.2%; Pred. No. 2,6e-162;  
Matches 756; Conservative 0; Mismatches 19; Indels 3; Gaps 3;  
QY 1049 AGAATGTGGCAGCAAGACATCCCCGAGCCAGAACTCCAGCTCAGGCTTACCAATGG 1108  
| | | | |  
Db 4 AGAATGTGGCAGCAAGACATCCCCGAGCCAGAACTCCAGCTCAGGCTTACCAATGG 63  
QY 1109 AAGTGGCCAGCCAGCCTTGGAGGGAAGATATCATCTGCTCCCTACAGGAGTG 1168  
| | | | |  
Db 64 AAGTGGCCAGCCAGCCTTGGAGGGAAGATATCATCTGCTCCCTACAGGAGTG 123  
QY 1169 GAAAAACCAAGTGGCTGTTTACATTGCCAAGATCACTTAGACAGAAAGAAAGCAT 1228  
| | | | |  
Db 124 GAAAAACCAAGTGGCTGTTTACATTGCCAAGATCACTTAGACAGAAAGAAAGCAT 183  
QY 1229 CTGAGCCTGAAAAAGTTATAGTCTTGTCAATAAAGTACTGCTAGTTGAACAGCTCTTC 1288  
| | | | |  
Db 184 CTGAGCCTGAAAAAGTTATAGTCTTGTCAATAAAGTACTGCTAGTTGAACAGCTCTTC 243  
QY 1289 GCAAGAGTTCACCAACATTTTGGAGAAATGTATCGTGTATTGATTAAAGTGTGATA 1348  
| | | | |  
Db 244 GCAAGAGTTCACCAACATTTTGGAGAAATGTATCGTGTATTGATTAAAGTGTGATA 303  
QY 1349 CCCAAGTGAATATATATTTCCAGAAAGTGTCAAGTCCGTGATATTAATCAGTACAG 1408  
| | | | |  
Db 304 CCCAAGTGAATATATATTTCCAGAAAGTGTCAAGTCCGTGATATTAATCAGTACAG 363  
QY 1409 CTCAATCTTGAAGAACTCCCTCTTAACCTTGAAAAATGAGAAAGTGTGTTCAAT 1468  
| | | | |  
Db 364 CTCAATCTTGAAGAACTCCCTCTTAACCTTGAAAAATGAGAAAGTGTGTTCAAT 423  
QY 1469 TGTCAAGCTTTTCCCTCATTTATCATTTGATGAATGTCAACACCAAGAAAGCAGTGT 1528  
| | | | |  
Db 424 TGTCAAGCTTTTCCCTCATTTATCATTTGATGAATGTCAACACCAAGAAAGCAGTGT 483  
QY 1529 ATTAATAATCATGAGGATTAATTTGATGCAAGAGTTGAAAAACAATAGACTCAAGAAAG 1588  
| | | | |  
Db 484 ATTAATAATCATGAGGATTAATTTGATGCAAGAGTTGAAAAACAATAGACTCAAGAAAG 543  
QY 1589 AAAACAACCAAGTATTCCTTCCCTCAGATCTGGGACTAAACAGCTTCACTGTGTTG 1648  
| | | | |  
Db 544 AAAACAACCAAGTATTCCTTCCCTCAGATCTGGGACTAAACAGCTTCACTGTGTTG 603  
QY 1649 GAGGGCCACGAGCAAGCCAAAGCTGAAGACACATTTTAAACATATGTCATCTTG 1708  
| | | | |  
Db 604 GAGGGCCACGAGCAAGCCAAAGCTGAAGACACATTTTAAACATATGTCATCTTG 662  
QY 1709 ATGCATTTACTATTAAACTGTTAAGAAACCTTGATCAACTGAAAAACCAATACAGG 1768  
| | | | |  
Db 663 ATGCATTTACTATTAAACTGTTAAGAAACCTTGATCAACTGAAAAACCAATACAGG 722  
QY 1769 AGCCATGCAAGAAAGTTGCCATTTGCAGATGCAACAGAGAAATCAATTAAGAGAA 1826

Db 723 GGCATGCGAGAAG-TTGGCGCTGGCGATGCGACC-GAGAAGATCCCTTTAAGGGAAA 778

RESULT 12  
BF983236 729 bp mRNA linear EST 23-JAN-2001  
LOCUS 602305873F1 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:4397083 5',  
DEFINITION mRNA sequence.  
ACCESSION BF983236 GI:12386048  
VERSION BF983236  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.  
REFERENCE 1 (bases 1 to 729)  
NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10097 row: a column: 20  
High quality sequence stop: 665.

FEATURES  
source 1..729  
location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4397083"  
/tissue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC 88"  
/note="Organ: small\_intestine; Vector: pCMV-SPORT6;  
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;  
oligo-dt primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 20.4%; Score 687; DB 2; Length 729;  
Best Local Similarity 98.9%; Pred. No. 2.5e-156;  
Matches 723; Conservative 0; Mismatches 5; Indels 3; Gaps 3;

QY 586 GAGGAGGAACGTGTGACAAATTGAGACAGAAACCGGATTGCTGCGAGAAACAATGGA 645  
Db 1 GAGGAGGAACGTGTGACAAATTGAGACAGAAACCGGATTGCTGCGAGAAACAATGGA 60

QY 646 AATGAATCAGGTGTAGAGAGCTAATAAAAGATTGTGCAGAAAGAAAAGTTCTCT 705  
Db 61 AATGAATCAGGTGTAGAGAGCTAATAAAAGATTGTGCAGAAAGAAAAGTTCTCT 120

QY 706 GCATTTCTGAATGTTCTTCTGTCAAACAGAAACAATGAACCTGTCCAAAGATTAAACAGGC 765  
Db 121 GCATTTCTGAATGTTCTTCTGTCAAACAGAAACAATGAACCTGTCCAAAGATTAAACAGGC 180

QY 766 TCTGATTGCTCAGAAAGCAATGCAAGATTGAGATTATCAACAAGTTGATGTCCTCAA 825  
Db 181 TCTGATTGCTCAGAAAGCAATGCAAGATTGAGATTATCAACAAGTTGATGTCCTCAA 240

QY 826 GTGAAGAGCAACTTCTTCAACCAAGTTCAAGCCAAATCTGGAGAGAGGCTGGGGC 885  
Db 241 GTGAAGAGCAACTTCTTCAACCAAGTTCAAGCCAAATCTGGAGAGAGGCTGGGGC 300

QY 886 ATGAGAAATACTCATCAGAAATCATCTTTTGCAGATTCTTGTAGTTTCAGAAATCAGAC 945

Db 301 ATGAGAAATACTCATCAGAAATCATCTTTTGCAGATTCTTGTAGTTTCAGAAATCAGAC 360

QY 946 ACAAGTTTGGCAGAAAGAGTGTACGCTGCTTAGATGAAAGTCTTTGACATTAACAGCAAC 1005  
Db 361 ACAAGTTTGGCAGAAAGAGTGTACGCTGCTTAGATGAAAGTCTTTGACATTAACAGCAAC 420

QY 1006 ATGGCAGTGAATTCAGGACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGA 1065  
Db 421 ATGGCAGTGAATTCAGGACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGA 480

QY 1066 GCATCCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCC 1125  
Db 481 GCATCCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAG-C 539

QY 1126 TTGGAAGGGAAGAAATATCATCATCTGCTCCCTTACAGGAGTGAAGAAACAGAGTGCT 1185  
Db 540 TTGGAAGGGAAGAAATATCATCATCTG-CTCCCTACAGGAGTGAAGAAACAGAGTGCT 598

QY 1186 GTTTACATTGGCCAAAGATCACTTAGACAAAGAAAGAAAGCATCTGAGCCTGAGAAAGTT 1245  
Db 599 GTTTACATTGGCCAAAGATCACTTAGAC-AGAGAAAGAAAGCATCTGAGCCTGAGAAAGTT 657

QY 1246 ATAGTCTTGTCAATAGTAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAAACCA 1305  
Db 658 ATAGTCTTGTCAATAGTAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCCAACCA 717

QY 1306 TTTTGAAGAA 1316  
Db 718 TTTTGAAGAA 728

RESULT 13  
DR422160 690 bp mRNA linear EST 29-JUN-2005  
LOCUS nav08g05.y1 Human pterygium. Unnormalized (nav) Homo sapiens cDNA  
DEFINITION clone nav08g05 5', mRNA sequence.  
ACCESSION DR422160  
DR422160.1 GI:68324176  
EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.  
REFERENCE 1 (bases 1 to 690)  
Wistow,G., Jaworski,C., Aryankalayil-John,M., Rowsey,J.J., Cox,C.,  
Reid,T., Dushku,N. and Carper, D.  
NEIBank analysis of Human pterygium  
Unpublished (2005)  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 08 row: 9 column: 05  
Seq primer: Universal M13 Reverse.  
FEATURES  
source 1..690  
location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="nav08g05"  
/tissue\_type="Pterygium"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="Human pterygium. Unnormalized (nav)"  
/note="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted  
from 9 pooled human pterygia. A directionally cloned cDNA  
library in the pCMVSPORT6 vector (Invitrogen) was  
constructed at Bioserve Biotechnology (Laurel MD)







US-09-023-655-48

Query Match 38.0%; Score 1278.8; DB 3; Length 1392;  
Best Local Similarity 94.1%; Pred. No. 0;  
Matches 1311; Conservative 0; Mismatches 77; Indels 5; Gaps 2;

QY 1976 ATTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGT 2035  
DB 1 ATTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGT 60  
QY 2036 ATACTCATCTTGAACCTTTCTATATATGAGAGAAAGATAGAAGTTTGCACTCATAGAG 2095  
DB 61 ATACTCATCTTGAACCTTTCTATATATGAGAGAAAGATAGAAGTTTGCACTCATANNN 120  
QY 2096 ATGATAGTATGAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2155  
DB 121 NNT 180  
QY 2156 TAAAGAAACCTTGAACCTGATGAACAGATAGATTCTCATGACTTTATTTTGAAA 2215  
DB 181 TAAAGAAACCTTGAACCTGATGAACAGATAGATTCTCATGACTTTATTTTGAAA 240  
QY 2216 ACAATAAATGTTGAAAAGCGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCA 2275  
DB 241 ACAATAAATGTTGAAAAGCGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCA 300  
QY 2276 AATTAAGAAATACCATTAATGAGCAATATATAGACTGAGAAATCAGCAAGAAATAA 2335  
DB 301 AATTAAGAAATACCATTAATGAGCAATATATAGACTGAGAAATCAGCAAGAAATAA 360  
QY 2336 TCTTTACAAAAACAGACAGAGTGATGCGCTTTCCAGTGATTAAGAAATGAAA 2395  
DB 361 TCTTTACAAAAACAGACAGAGTGATGCGCTTTCCAGTGATTAAGAAATGAAA 420  
QY 2396 AATTGCTGAAGTAGAGTCAAAAGCCACCATCTGATTTGAGCTGACACAGCAGTGA 2455  
DB 421 AATTGCTGAAGTAGAGTCAAAAGCCACCATCTGATTTGAGCTGACACAGCAGTGA 480  
QY 2456 TCAAAACCCATGACACAGAGTGAACAAAAAGAGTCAATTAATTTCCGACTGGA 2515  
DB 481 TCAAAACCCATGACACAGAGTGAACAAAAAGAGTCAATTAATTTCCGACTGGA 540  
QY 2516 TCAATCTGCTTATCGCTACACAGTGGCAGAGAAAGTCTGATATTTAAAGATGTACA 2575  
DB 541 TAAATCTGCTTATCGCTACACAGTGGCAGAGAAAGTCTGATATTTAAAGATGTACA 600  
QY 2576 TTGTTATCCGTTATGCTGCTACCAATGAATAGCCATGTCACAGGCCGCTGTCGAG 2635  
DB 601 TTGTTATCCGTTATGCTGCTACCAATGAATAGCCATGTCACAGGCCGCTGTCGAG 660  
QY 2636 CCAGAGCTGATGAGAGCACTACGTCCTGTTGCTCACAGTGTTCAGAGTTATCGAAC 2695  
DB 661 CCAGAGCTGATGAGAGCACTACGTCCTGTTGCTCACAGTGTTCAGAGTTATCGAAC 720  
QY 2696 ATGAGACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAGCTATATCTGTCTCAA 2755  
DB 721 GTGAGACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAGCTATATCTGTCTCAA 780  
QY 2756 ATATGAAACCCAGAGAGTATGCTCATAGAATTTTGAATTTACAGATGCAAAAGTATATGG 2815  
DB 781 ATATGAAACCCAGAGAGTATGCTCATAGAATTTTGAATTTACAGATGCAAAAGTATATGG 840  
QY 2816 AAAAGAAATGAAAACCAAGAAATATTTGCCAAGCATTAACAAGATTAACCATCACTAA 2875  
DB 841 AAAAGAAATGAAAACCAAGAAATATTTGCCAAGCATTAACAAGATTAACCATCACTAA 900  
QY 2876 TAACTTTCCTTTGCAAAAACCTGACGTGTGCTGCTGTTCTGGGAAGATATCCATGTAA 2935  
DB 901 TAACTTTCCTTTGCAAAAACCTGACGTGTGCTGCTGTTCTGGGAAGATATCCATGTAA 960  
QY 2936 TTGAGAAAATGCATCAGTCAATATGACCCCAAGAAATTCAGGAACCTTACATTTGTAAGAG 2995  
DB 961 TTGAGAAAATGCATCAGTCAATATGACCCCAAGAAATTCAGGAACCTTACATTTGTAAGAG 1020

QY 2996 AAAACAAAGCACTGCAAAAAGAGTGTGCGGCACTATCAAAATAATGSTGAATCATCTGCA 3055  
DB 1021 AAAACAAAGCACTGCAAAAAGAGTGTGCGGCACTATCAAAATAATGSTGAATCATCTGCA 1080  
QY 3056 AATGTGGCCAGGCTTTGGGGAACAATGATGTGTCACAAAGGCTTAGATTGCTGTCTCA 3115  
DB 1081 AATGTGGCCAGGCTTTGGGGAACAATGATGTGTCACAAAGGCTTAGATTGCTGTCTCA 1140  
QY 3116 AAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCAACAAAGAAACAATACAAAAAGT 3175  
DB 1141 AAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCAACAAAGAAACAATACAAAAAGT 1200  
QY 3176 GGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTGAGATGCTGTTATTAGTG 3235  
DB 1201 GGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTGAGATGCTGTTATTAGTG 1260  
QY 3236 ATGAGATTAGCACTGATTGAAGATTTCTTTAAATACATATCATTAACATTAATAT 3295  
DB 1261 ATGAGATTAGCACTGATTGAAGATTTCTTTAAATACATATCATTAACATTAATAT 1318  
QY 3296 GATTATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 3352  
DB 1319 TGTATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1378  
QY 3353 TTGTTTACTCTG 3365  
DB 1379 TTGTTTACTCTG 1391

RESULT 2  
US-09-907-907A-32  
: Sequence 32, Application US/09907907A  
: Patent No. 6951923

: GENERAL INFORMATION:  
: APPLICANT: Fisher, Paul B.  
: APPLICANT: Leszczynska, Magdalena  
: TITLE OF INVENTION: GENES DISPLAYING ENHANCED EXPRESSION DURING CELLULAR SENESENCE A  
: FILE REFERENCE: A34584-A-PCT-USA (070050.1664)  
: CURRENT APPLICATION NUMBER: US/09/907,907A  
: CURRENT FILING DATE: 2001-07-16  
: PRIOR APPLICATION NUMBER: US 09/243,277  
: PRIOR FILING DATE: 1999-02-02  
: NUMBER OF SEQ ID NOS: 51  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 32  
: LENGTH: 458  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: unsure  
: LOCATION: 342, 355, 365, 368, 375, 381, 385, 414, 445  
: OTHER INFORMATION: a or c or g or t  
US-09-907-907A-32

Query Match 11.2%; Score 376.4; DB 3; Length 458;  
Best Local Similarity 97.7%; Pred. No. 6.6e-91;  
Matches 388; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 2969 AATTCAGGAACCTTACATTGTGAAGAGAAACAAAGCACTGCAAAAGAGTGTGCCGACT 3028  
DB 1 AATTCAGGAACCTTACATTGTGAAGAGAAACAAAGCACTGCAAAAGAGTGTGCCGACT 60  
QY 3029 ATCAAAATAATGTTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGC 3088  
DB 61 ATCAAAATAATGTTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGC 120  
QY 3089 ACAAGGCTTAGATTGCTGCTCAAAATAAGGAATTTTGTAGTGTCTTCAAAAATA 3148  
DB 121 ACAAGGCTTAGATTGCTGCTCAAAATAAGGAATTTTGTAGTGTCTTCAAAAATA 180  
QY 3149 ATTCAACAAGAAACAATAACAAGAGTGGTAGAATTACCTATCATTTCCCAATCTTG 3208

Db 181 ATCAACAAGAAACATACAAAAAGTGGTAGAATTAACCTATACATTCCCAATCTTG 240

QY 3209 ACTATTCAGAAATGCTGTTTATTTAGTAGAGATTAGCACTTGATTGAGATTCTTTA 3268

Db 241 ACTATTCAGAAATGCTGTTTATTTAGTAGAGATTAGCACTTGATTGAGATTCTTTA 300

QY 3269 AAATACTATCAGTTAAACATTTAATATGATTATGATTATGATTATGCTACAGAA 3328

Db 301 AAATACTATCAGTTAAACATTTAATATGATTATGATTATGATTATGCTACAG-A 359

QY 3329 CTGACATAGAATCAATTAATGATTGTTTACTCTG 3365

Db 360 CTGACNTANGAATCANTAAANGATNGTTTACTCTG 396

## RESULT 3

US-08-143-576-7

; Sequence 7, Application US/08143576

; Patent No. 5643761

; GENERAL INFORMATION:

; APPLICANT: Fisher, Paul B.

; APPLICANT: Jiang, Hongping

; TITLE OF INVENTION: METHOD FOR GENERATING A SUBTRACTED

; TITLE OF INVENTION: CDNA LIBRARY AND USES OF THE GENERATED LIBRARY

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: John P. White, c/o Cooper & Dunham

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/143,576

; FILING DATE: 25-OCT-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 43563/JPW/AKC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 977-9550

; TELEFAX: (212) 664-0525

; TELEX: 422523 COOP UI

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 301 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-08-143-576-7

Query Match 8.5%; Score 285.4; DB 2; Length 301;

Best Local Similarity 99.3%; Pred. No. 1.6e-66;

Matches 297; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1592 ACAAACCAAGTATCCCTTCTCAGATATCTGGACTAACAGCTTCACTGGTGTGGAG 1651

Db 1 ACAAACCAAGTATCCCTTCTCAGATATCTGGACTAACAGCTTCACTGGTGTGGAG 60

QY 1652 GGGCCACGAAGCAAGCCAAAGCTGAAGACACATTTTAAACTATGTGCCAATCTTGATG 1711

Db 61 GGGCCACGAAGCAAGCCAAAGCTGAAGACACATTTTAAACTATGTGTGCTATCTTGATG 120

QY 1712 CATTACTATTAACCTGTTAAAGAAAACCTTGATCACTGAAAACCAATACAGAGC 1771

Db 121 CATTACTATTAACCTGTTAAAGAAAACCTTGATCACTGAAAACCAATACAGAG- 179

QY 1772 CATGCAAGAAGTTGGCATTGCGATGCAACCAAGAGATCCATTAAAGAGAACTTC 1831

Db 180 CATGCAAGAAGTTGGCATTGCGATGCAACCAAGAGATCCATTAAAGAGAACTTC 239

QY 1832 TAGAATAATGACAGAGATTCAAACTTATGTCAATAGTCCAAATGTCAGATTTGGA 1890

Db 240 TAGAATAATGACAGAGATTCAAACTTATGTCAATAGTCCAAATGTCAGATTTGGA 298

## RESULT 4

US-09-221-268D-12

; Sequence 12, Application US/09221268D

; Patent No. 6720408

; GENERAL INFORMATION:

; APPLICANT: Fisher, Paul B.

; APPLICANT: Jiang, Hongping

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH SPECIFICALLY HYBRIDIZE WITH

; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE ENCODING MDA-7 AND PHARMACEUTICAL COMPOSIT

; FILE REFERENCE: A34534-A-A (070050.1637)

; CURRENT APPLICATION NUMBER: US/09/221,268D

; PRIOR FILING DATE: 1998-12-23

; PRIOR APPLICATION NUMBER: 08/316,537

; PRIOR FILING DATE: 1994-09-30

; PRIOR APPLICATION NUMBER: 08/143,576

; PRIOR FILING DATE: 1993-10-27

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO: 12

; LENGTH: 301

; TYPE: DNA

; ORGANISM: homo sapiens

; US-09-221-268D-12

Query Match 8.5%; Score 285.4; DB 3; Length 301;

Best Local Similarity 99.3%; Pred. No. 1.6e-66;

Matches 297; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1592 ACAAACCAAGTATCCCTTCTCAGATATCTGGACTAACAGCTTCACTGGTGTGGAG 1651

Db 1 ACAAACCAAGTATCCCTTCTCAGATATCTGGACTAACAGCTTCACTGGTGTGGAG 60

QY 1652 GGGCCACGAAGCAAGCCAAAGCTGAAGACACATTTTAAACTATGTGCCAATCTTGATG 1711

Db 61 GGGCCACGAAGCAAGCCAAAGCTGAAGACACATTTTAAACTATGTGCTATCTTGATG 120

QY 1712 CATTACTATTAACCTGTTAAAGAAAACCTTGATCACTGAAAACCAATACAGAGC 1771

Db 121 CATTACTATTAACCTGTTAAAGAAAACCTTGATCACTGAAAACCAATACAGAG- 179

QY 1772 CATGCAAGAAGTTGGCATTGCGATGCAACCAAGAGATCCATTAAAGAGAACTTC 1831

Db 180 CATGCAAGAAGTTGGCATTGCGATGCAACCAAGAGATCCATTAAAGAGAACTTC 239

QY 1832 TAGAATAATGACAGAGATTCAAACTTATGTCAATAGTCCAAATGTCAGATTTGGA 1890

Db 240 TAGAATAATGACAGAGATTCAAACTTATGTCAATAGTCCAAATGTCAGATTTGGA 298

## RESULT 5

US-09-799-451-771

; Sequence 771, Application US/09799451

; Patent No. 6783969

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

; APPLICANT: Ren, Feiyan

; APPLICANT: Zhang, Jie

APPLICANT: Xue, Aidong J.  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Ma, Yuning  
APPLICANT: Yamazaki, Victoria  
APPLICANT: Chen, Rui-hong  
APPLICANT: Wang, Zhiwei  
APPLICANT: Wang, Dunrui  
APPLICANT: Yang, Yonghong  
APPLICANT: Wehrman, Tom  
APPLICANT: Ghosh, Reena  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 803  
CURRENT APPLICATION NUMBER: US/09/799,451  
CURRENT FILING DATE: 2001-03-05  
NUMBER OF SEQ ID NOS: 948  
SOFTWARE: pt\_fl\_genes Version 2.0  
SEQ ID NO 771  
LENGTH: 2590  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (190)..(2223)  
US-09-799-451-771

Query Match 5.3%; Score 179.6; DB 3; Length 2590;  
Best Local Similarity 47.6%; Pred. No. 1.6e-37;  
Matches 999; Conservative 0; Mismatches 949; Indels 150; Gaps 10;

QY 1048 GAGAATGTGGCAGCAAGACATCCCGGAGCCAGACTCCAGCTCAGGCTTACCAATG 1107  
DB 154 GGGACTGCCCCGGCAGACCTACTAGAGCAGATGAGCTTCGGTCTCAATG 213  
QY 1108 GAAGTGGCCAGCCAGCCTTGAAGGGAAGATATCATCTGCTCCCTACAGGAGT 1167  
DB 214 GAGGTGATCATGCTGCTGCTGGAGGGCAAGATATCATCTGCTGCTCCACGGGTCC 273  
QY 1168 GGAAGAACCAAGAGTGGCTGTTTACATTGCCAAGATCACTTAGACAAGAGAAAAAGCA 1227  
DB 274 GGAAGACCCGGGGCGCTGTTATGTGGCCAAAGCGGCACTAGAGACTGTGATGAGC- 332  
QY 1228 TCTGAGCCTGAAAAAGTTATAGTCTTGTCAATTAAGTACTGCTAGTTGAACAGCTTTC 1287  
DB 333 -----CAAGTGGTTGTATTGTCACACAGGCTGACCTGTGACCCAGC--AT 378  
QY 1288 CGCAAGAGTTCCAACCATTTTGAAGAAATGTAATGTAATGTAATGTAATGTAATGTAAT 1347  
DB 379 GGTGAAGAGTTCAAGCGCATGCTGATGAGCGTGACCGTGACCAACCTGATGGGAGC 438  
QY 1348 ACCCACTGAAAAATATCATTTCCAGAAGTGTCAAGTCTGTGATATATATATATATATAT 1407  
DB 439 ATGGGACCAAGTGTGCTGGCTTGGCCACCTGGCCCGGTGCCATGACCTGCTCATCTGCACA 498  
QY 1408 GCTCAATTCCTGAAAACTCCTCTTAACCTTGAATAATGGAAGAGATGCTGCTTCAA 1467  
DB 499 GCAGAGCTTCTGAGATGCACTGACCGCCCGA-----GGAGAGAGGACGCTGGAG 552  
QY 1468 TTGTCAGACTTTTCCCTCATTTATCATGTAGATGATGATGATGATGATGATGATGATGAT 1527  
DB 553 CTCACTGTCTTCTCCCTGATCGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 612  
QY 1528 TATTAATAATCATGAGGCAATATTGATGAGAGTTGAAAAACAATAGACTCAAGAAA 1587  
DB 613 TACAACGTATCATGAGCCAGTACT-----AGACTTAACTCCAGAGG 657  
QY 1588 GAAAAACAACCAAGTATTCCTCTTCTCAGATATCTGGACTTAACAGCTTCACTGGTGT 1647  
DB- 658 GCACA-----GCCGCTAACCCCAAGGTGCTGGGTCTCACAGCCTCCCGCAGGCACT 705  
QY 1648 GGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTT 1707

DB 706 GGGGGGGCTCCAAACTCGATGGGGCCATCAACCAAGCTCCTGACGCTCTGTGCCAATCTG 765  
QY 1708 GATGCATTACTATTAAACTGTTAAAGAAAACTTGATCAACTGAAGAAAAACCAATACAG 1767  
DB 766 GACACGTGTGATCATATGTCAACCCAGAACTGCTGCCCCAGCTGACAGACACAGCCAA 825  
QY 1768 GAGCCATGCAAGAGTTTGGCCATGTCAGATGCAACGAGAGAGATCAATTAAGAGAAA 1827  
DB 826 CAGCCTTGCAAAACAGTAAACCTCTGCCACAGGCGCAGCCAGATCGTTGGGACTTG 885  
QY 1828 CTTCTAGAATAATGACAAGATTCAACTTAATGTCAAATGAGTCAATGTCA--GAT 1884  
DB 886 CTGAAGAAGCTCATGACCAATCCATGACCACTGAGATGCTGAGCTTGAGCCGAAA 945  
QY 1885 TTGGAATCAACCTTATGAACAATGGGCCATTCAAATGAAAAAAAGCTGCAAAAAA 1944  
DB 946 TTGGGACGCAAAATGATGAGCAGCAGAGGTGTAAGCTGAGTGAAGCTGCGGCTTGGCT 1005  
QY 1945 GGAATCGCAAGAGCTGTTTGTGCAAGACATTTGAGGAAGTACATGAGGCCCTACAA 2004  
DB 1006 GGGCTTACAGACCAAGGGGTATGCGCTTCACTGAGCGGCTACATGACGCGTCTC 1065  
QY 2005 ATTAATGACACAATTTGAATGATAGATGCTATCTCATCTTGAACTTCTATATGAA 2064  
DB 1066 ATCCATGACACCGTCCGCGCGCTGATGCTTGGCTGCGCTGACAGATTTCTATCACAG 1125  
QY 2065 GAGAAAGATAAGAGTTTGCAGTCAATAGAGATGATAGTGAAGGTGTGATGATGAG 2124  
DB 1126 GAGCAGCTCACTAAACCAAGATCCTGTGTG----- 1156  
QY 2125 TATTTGATGTGATGATGAGATGAGATGATTTAAAGAAACCTTGAAACTGATGAACA 2184  
DB 1157 -----CC 1158  
QY 2185 GATAGATTTCTCATGACTTATTTTGAACAATAAATGTTGAAGAGCTGGCTGAA 2244  
DB 1159 GAGCGCGGCTGCTGCTGCTGCTGCTGATGACCGCAAGATGAGCTGGCCACTTGGC--A 1215  
QY 2245 AACCCAGATATGAAAAAGTGAACCAATTAAGAAATACATTAATGAGCAATAT 2304  
DB 1216 ACTCATGGCCAGAGATCCAAACTGAGATGCTGAAAAAGATCTGCAAAAGGCAATTTC 1275  
QY 2305 ACTAGACTGAGAAATCAGCAGAGAAATATCTTTACAAAAACAGACAGATGCATAT 2364  
DB 1276 AGTAGCTCT--AACAGCCTCGGGTATCATCTTCAACCGCACCGCCAAAGCGCACAC 1332  
QY 2365 GCGCTTCCAGTGAATTAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 2424  
DB 1333 TCCCTCCTGCTGTGGCTCCAGCAGCAGAGGCGCTGCAAGCTGTGACATCCGGGCCAG 1392  
QY 2425 CATCTGATTGAGCTGACACAGCAGTGAATTCACCAACCATGACACAGATGAACAAAA 2484  
DB 1393 CTAAGTATGGGGCTGGGAAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1452  
QY 2485 GAAGTCAATTAATTTGCACTGGAATAATCAATCTGTTATCGCTAACACAGTGGCA 2544  
DB 1453 GAAGTATCCAGAAAGTTCCAAGATGAACCTTCTGTTGTCGCAAGTGTGGCG 1512  
QY 2545 GAAAGAGTCTGATATTAAAGATGAACATTTATCCGTTATGCTCTGTCACCAAT 2604  
DB 1513 GAGAGGGGCTGACATCCCAATGCAATGTGTGTGCTGCTTATGGCTCTTGACCAAT 1572  
QY 2605 GAAATAGCCATGCTCAGGCGCGGTGTCAGCAGCAGAGCTGATGAGACCACTAGCTCTG 2664  
DB 1573 GAAATCTCATGTGTCCAGGCCAGGGGCGGTGCTCCGCGGATCAGAGTATATACGCGTTT 1632  
QY 2665 GTTGCTCACAGTGTTCAGGAGTTATCGAATGAGACAGATTATGATTTCCGAGAGAG 2724  
DB 1633 GTAGCAACTGAAGGTACCGGAGCTGAAGCGGAGCTGATCAACGAGCGCTGAGAGAG 1692  
QY 2725 ATGATGTATAAGCTATATATGTTGTTCAAAATATGAACCAAGAGGATATGCTCATAG 2784

Db 1693 CTGATGAGACGAGCTGGCTGCTGTGACAGAAATGACCAGCCGAGTACAGGCCAAG 1752  
QY 2785 ATTTGGAATTACAGATGCAAGATATAATGAAAAAATGAAACCAAGAGAAATATT 2844  
Db 1753 ATCCGGATCTGACAGCAGCAGCCTTGACCAAGCGCGCCAGGACGCCAGCGGAG 1812  
QY 2845 GCCAAGCATTAAGAATAACCCATCACTTAATTACTTCCCTTGCAAAAACAGCAGTGTG 2904  
Db 1813 AACCAAGCGCAGCAGTCCCAAGTGAGCAGCTGCACTCTGCACTCAATGATGTGTG 1872  
QY 2905 CTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAGAAATGCATCAGTCAATATGACC 2964  
Db 1873 GCTGTGGGCCATGCGACGACCTGCGGAGGTGAGGGCACCCAACATGTCAATGTGAAC 1932  
QY 2965 CCAGAAATCAAGAACTTACATTTGTAAGAGAAAAACAAGCACTGCAAAAAGAGTGTGC 3024  
Db 1933 CCCAAGCTTCTGAAGTAACTATTAATGTCAGGAGATCCTGTGTCATCAACAAAGTCTTC 1992  
QY 3025 GACTATCAATAATGATGTAATCATCTGC-----AAATGCGCCAGGCTTGGGGAACA 3078  
Db 1993 AAGACTGGAAGCCTGGGGGTGTCTCATCACTGACAGAACTGTGGGAGGTCTGGGCTTG 2052  
QY 3079 ATGATGTCGACAAAGGCTTAGATTGCTGCTCAAAATAAGAAATTTTGTAGTGG 3136  
Db 2053 CAGATGATCTACAGTCACTGAGAGCTGCCAGTGTCTCAAAGTCCGACATGTCTGTGG 2110

RESULT 6

US-08-143-576-6  
; Sequence 6, Application US/08143576  
; Patent No. 5643761

GENERAL INFORMATION:

APPLICANT: Fisher, Paul B.  
APPLICANT: Jiang, Hongping  
TITLE OF INVENTION: METHOD FOR GENERATING A SUBTRACTED  
TITLE OF INVENTION: CDNA LIBRARY AND USES OF THE GENERATED LIBRARY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White, c/o Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/143,576  
FILING DATE: 25-OCT-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 43563/JPW/AKC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 664-0525

TELEX: 422523 COOP UI

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 193 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-143-576-6

Query Match 3.7%; Score 124.6; DB 2; Length 193;

Best Local Similarity 96.5%; Pred. No. 2.3e-23;  
Matches 138; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
QY 3007 CTGCAAAAGAGTGTGCCGACTATCAATAATGTGAATCATCTGCAATGTGCCAG 3066  
Db 1 CTGCAAAAGAGTGTGCCGACTAT-CAATAATGTGAATCATCTGCAATGTGCCAG 59  
QY 3067 GCTTGGGAACAATGATGTGACAAAGGCTTAGATTGCTGCTCAAAATAGGAAT 3126  
Db 60 GCTTGGGAACAATGATGTGACAAAGGCTTAGATTGCTGCTCAAAATAGGAAT 119  
QY 3127 TTTGTAGTGTTCATAAATAA 3149  
Db 120 TTTGTAGTGTTCATAAATAATCA 142

RESULT 7

US-09-221-268D-11  
; Sequence 11, Application US/09221268D  
; Patent No. 6720408

GENERAL INFORMATION:

APPLICANT: Fisher, Paul B.

APPLICANT: Jiang, Hongping

TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH SPECIFICALLY HYBRIDIZE WITH

TITLE OF INVENTION: NUCLEIC ACID SEQUENCE ENCODING MDA-7 AND PHARMACEUTICAL COMPOSIT

FILE REFERENCE: A34534-A-A (070050.1637)

CURRENT APPLICATION NUMBER: US/09/221,268D

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 08/316,537

PRIOR FILING DATE: 1994-09-30

PRIOR APPLICATION NUMBER: 08/143,576

PRIOR FILING DATE: 1993-10-27

NUMBER OF SEQ ID NOS: 30

SOFTWARE: FastSeq for windows Version 4.0

SEQ ID NO 11

LENGTH: 193

TYPE: DNA

ORGANISM: homo sapiens

US-09-221-268D-11

Query Match 3.7%; Score 124.6; DB 3; Length 193;  
Best Local Similarity 96.5%; Pred. No. 2.3e-23;  
Matches 138; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 3007 CTGCAAAAGAGTGTGCCGACTATCAATAATGTGAATCATCTGCAATGTGCCAG 3066

Db 1 CTGCAAAAGAGTGTGCCGACTAT-CAATAATGTGAATCATCTGCAATGTGCCAG 59

QY 3067 GCTTGGGAACAATGATGTGACAAAGGCTTAGATTGCTGCTCAAAATAGGAAT 3126

Db 60 GCTTGGGAACAATGATGTGACAAAGGCTTAGATTGCTGCTCAAAATAGGAAT 119

QY 3127 TTTGTAGTGTTCATAAATAA 3149

Db 120 TTTGTAGTGTTCATAAATAATCA 142

US-10-131-827-8333

; Sequence 8333, Application US/10131827

; Patent No. 6905827

GENERAL INFORMATION:

APPLICANT: Wohlgenuth, Jay

APPLICANT: Fry, Kirk

APPLICANT: Woodward, Robert

APPLICANT: Ly, Ngoc

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE

TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES

FILE REFERENCE: 506612000120

CURRENT APPLICATION NUMBER: US/10/131,827

CURRENT FILING DATE: 2002-09-06

PRIOR APPLICATION NUMBER: US 10/006,290

PRIOR FILING DATE: 2001-10-22

US-10-131-827-8333

Query Match 3.7%; Score 124.6; DB 3; Length 193;

```
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8333
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-131-827-8333

Query Match      2.7%; Score 91; DB 3; Length 530;
Best Local Similarity 57.3%; Pred. No. 4.6e-14;
Matches 279; Conservative 0; Mismatches 200; Indels 8; Gaps 6;

QY 2024 TGATAGATGCGTATCTCATCTTGAACTTCTATATGAAGAGAAAGTAAAGTTTG 2083
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 TTATATATGCGTATAGTCATCTATCGTATATCTATATTAATTATTAATATGTGTTATCTG 97

QY 2084 CAGTCATAGAGATGATAGTAGAGGGTGTGATGATGATGATTTGTA-TGGTGAAGAA 2142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 CAGTGAATAATTATGATCTTGATGTGCGTGTGATGATTAATTGTGACTCGTCATTAC 157

QY 2143 GATGAGATGAT-TTAAAGAAACCTTGAACCTGATGAAGACAGATAGATTTCATGAC 2201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 TATGAAGATGATCTTGAGGCTGTGATCTGTGATGATGAAGGTGATGAGTGTGATGCA 217

QY 2202 TTTATTTT--TGAACAATAAATGTTGAAAAAGGCTGGCTGAAAAACCCAGATATGAA 2259
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218 CAGGATTTTAAAGATGGCTCTAAAGTGTGCATAGAGGTTGGTGTCAACAATAATAGAA 277

QY 2260 AATGA-AAAGCTGACCAATTAGAAATACCATAATGAGCAATATACTAGGACTGAGG 2317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 GATGAGCTAGGTGATCGGATGATGAAATACACTTGTGAGCGGTGATTCTAAGACTGAA 337

QY 2318 AATCAGCAGCAGAAATAATCTTTACAAAAACGACAGAGTGCATATGCGTTTCCAGT 2377
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 ACTC-GCTTGATGATATCCTCTAAGATGAACGAGTGCATGCTGATCTCCGCTTCTCAG 396

QY 2378 GGATTACTGA-AAATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGA 2436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 397 GGATCACTGAGCGATGAAGACAGATGAGGTATGAGTGTGCTCGACGCGCGGTGCT 456

QY 2437 GCTGACACAGCAGTGAAGTCAAAACCATGACACAGATGAACAAAAGATCATTAGT 2496
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 457 GCTGTGCGCCGTTGGCGATGTGACAGCCTGAGAGATGATATGTGCATAACGAGATCCAGATG 516

QY 2497 AAATTTC 2503
   ||| |||
Db 517 ACAATTTC 523
```

```
RESULT 9
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
; US-08-232-463-14
```

```
Query Match      2.1%; Score 69.4; DB 2; Length 7218;
Best Local Similarity 4.8%; Pred. No. 1.3e-07;
Matches 19; Conservative 230; Mismatches 146; Indels 0; Gaps 0;
```

```
QY 1876 ATGTCAATTTTGAACCTCAACCTATGAACAATGGCCATTCAATGAAAAAAGCT 1935
   || ||| ||| ||| ||| : : : : : : : : : : : : : : : : : : : : : :
Db 1451 ATAGAAGATTGTGACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1392

QY 1936 GCAAAAAAGAAATCGCAAGACGTGTTGTGACAGAACATTTGAGGAAGTACAAATGAG 1995
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1391 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1332

QY 1996 GCCCTACAATTAATGACACAATTCGAATGATAGATGCGTATCTCATCTTGAACCTTC 2055
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1331 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1272

QY 2056 TATTAATGAAGAAAGATTAAGAGTTTGCACTCATAGAAAGATGATGATGAGGCTGT 2115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1271 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1212

QY 2116 GATGATGAGTATTTGTGATGCTGATGAAGATGAGATGATTAAGAAACCTTTGAAACTG 2175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1211 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1152

QY 2176 GATGAACAGATAGATTTCTCATGACTTATTTTGAACAATAAATGTTGAAAGG 2235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1151 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1092

QY 2236 CTGCTGAAAAACCAAGATATGAATAAGCT 2270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1091 RRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCT 1057
```

```
RESULT 10
US-09-949-016-17284
; Sequence 17284, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17284  
; LENGTH: 12214  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17284

Query Match 2.0%; Score 66.6; DB 3; Length 12214;  
Best Local Similarity 62.1%; Pred. No. 1e-06;  
Matches 105; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

OY 1078 CCAGAACTCCAGCTCAGGCTTACCAATGGAAGTGGCCAGCCGCTTGAAAGGAG 1137  
DB 11430 CCCAGAAATGAGCTTCGGTCTCTACCAATGGAGGTGATCATGCTGCCCTGAGGGCAG 11489  
OY 1138 AATATCATCATCTGCTCTCCCTACAGGAGTGAAGAAACAGAGTGGCTTTACATTGCC 1197  
DB 11490 AATATCATCATCTGCTGCTCCACGGGTGCCGGAAGACCCGGGCGCTGTATGTGCC 11549  
OY 1198 AAGATCACTTAGACAAGAAAAAGCATCTGAGCCTGGAAGTTA 1246  
DB 11550 AAGCGGACCTAGAGACTGTGATGAGCCAGAGTGTGTATGTGTC 11598

RESULT 11

US-09-590-968B-1  
; Sequence 1, Application US/09590968B  
; Patent No. 6737561  
; GENERAL INFORMATION:  
; APPLICANT: Ray, Animesh  
; APPLICANT: Golden, Teresa Ann  
; TITLE OF INVENTION: GENE ENCODING SHORT INTEGUMENTS AND USES THEREOF  
; FILE REFERENCE: 176/60581  
; CURRENT APPLICATION NUMBER: US/09/590,968B  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/138,316  
; PRIOR FILING DATE: 1999-06-09  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 6184  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-590-968B-1

Query Match 1.9%; Score 63; DB 3; Length 6184;  
Best Local Similarity 54.5%; Pred. No. 6.4e-06;  
Matches 126; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

OY 2464 ATGACACAGAATGAACAAAAAGAGTCATTAGTAATTTGCACTGGAATAATCAATCTG 2523  
DB 2492 ATGAATCATCTCAATGACAGATACAAATTCGAATTCGAGATGGGCATGTGACACTG 2551  
OY 2524 CTTATCGTACCAAGTGGCAGAGAGGTCTGATATTAAAGATGTAACATTGTTATC 2583  
DB 2552 TTAGTGGCAACAGCGTCTGAGGAAGACTTGATATTAGGCAATGTAACGTTGTTATG 2611  
OY 2584 CGTTATGCTCTGTCACCAATGAATAGCCATGCTCCAGGCCCCGTGTCGAGCCAGAGCT 2643  
DB 2612 CGTTTCGACCTTGCAAGACGGTGTGTCGATACATTCACTCTGTCGCGGCAAGAAAG 2671  
OY 2644 GATGAGAGCACTAGCTCTGCTGTTGCTCAGAGTGTTCAGAGTTATCGAA 2694  
DB 2672 CCTGATCAGACTACATACTCATGTGTGAGAGAGAAATGTATCTCAAGCA 2722

RESULT 12  
US-09-853-768-10

; Sequence 10, Application US/09853768  
; Patent No. 6444466  
; GENERAL INFORMATION:  
; APPLICANT: Donna T. Ward  
; APPLICANT: Andrew T. Walt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF HELICASE-MOI EXPRESSION  
; FILE REFERENCE: RTS-0217  
; CURRENT APPLICATION NUMBER: US/09/853,768  
; PRIOR FILING DATE: 2001-05-10  
; NUMBER OF SEQ ID NOS: 91  
; SEQ ID NO 10  
; LENGTH: 5852  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(4568)  
US-09-853-768-10

Query Match 1.8%; Score 62; DB 3; Length 5852;  
Best Local Similarity 60.8%; Pred. No. 1.2e-05;  
Matches 101; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

OY 2477 AACAAAAGAGTCAATTAGTAATTTGCGCACTGGAATAATCAATCTGCTTATCGCTACCA 2536  
DB 298 AACAGAGAGGTACTTAGGAATTTGAGCACATGAGACCAACTGCTTATTGCAACAA 357  
OY 2537 CAGTGCAGAGAGGCTCTGATATTAAAGATGTAACATTGTTATCCGTTATGCTTCG 2596  
DB 358 GTATTGTAGAGAGGCTGTGATATTACCAAAATGCACACTTGGTGTGTTGATTGTC 417  
OY 2597 TCACCAATGAATAGCCATGCTCCAGGCCCTGTCGAGCCAGAGC 2642  
DB 418 CCACAGAAATGATCTCTATGTTCAATCTAAAGAGAGCAAGGCGC 463

RESULT 13

US-09-853-768-3  
; Sequence 3, Application US/09853768  
; Patent No. 6444466  
; GENERAL INFORMATION:  
; APPLICANT: Donna T. Ward  
; APPLICANT: Andrew T. Walt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF HELICASE-MOI EXPRESSION  
; FILE REFERENCE: RTS-0217  
; CURRENT APPLICATION NUMBER: US/09/853,768  
; PRIOR FILING DATE: 2001-05-10  
; NUMBER OF SEQ ID NOS: 91  
; SEQ ID NO 3  
; LENGTH: 7037  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (183)...(5957)  
US-09-853-768-3

Query Match 1.8%; Score 62; DB 3; Length 7037;  
Best Local Similarity 60.8%; Pred. No. 1.3e-05;  
Matches 101; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

OY 2477 AACAAAAGAGTCAATTAGTAATTTGCGCACTGGAATAATCAATCTGCTTATCGCTACCA 2536  
DB 1687 AACAGAGAGGTACTTAGGAATTTGAGCACATGAGACCAACTGCTTATTGCAACAA 1746  
OY 2537 CAGTGCAGAGAGGCTCTGATATTAAAGATGTAACATTGTTATCCGTTATGCTTCG 2596  
DB 1747 GTATTGTAGAGAGGCTGTGATATTACAAATGCAACTTGGTGTGTTGATTGTCG 1806  
OY 2597 TCACCAATGAATAGCCATGCTCCAGGCCCTGTCGAGCCAGAGC 2642  
DB 1807 CCACAGAAATGATCTCTATGTTCAATCTAAAGAGAGCAAGGCGC 1852



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: March 13, 2006, 15:10:59 ; Search time 12661 Seconds  
(without alignments)  
4601.891 Million cell updates/sec

Title: US-09-515-363C-2  
Perfect score: 5311  
Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNLDYSECLFSDED 1025

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5883141 segs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEl=frame+\_p2n.model -DEV=xlh  
-Q=/abs/ABSSWEB.spool/US09515363/runat\_10032006\_181128\_7937/app.query.fasta\_1  
-DB=GenEmbl -QFWT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOFC=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdf -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs04  
-USER=US09515363 @CCN\_1\_1\_4939 @runat\_10032006\_181128\_7937 -NCPU=6 -ICPU=3  
-NO\_MMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_in: \*  
3: gb\_env: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pr: \*  
9: gb\_ro: \*  
10: gb\_sts: \*  
11: gb\_sy: \*  
12: gb\_un: \*  
13: gb\_vi: \*  
14: gb\_htg: \*  
15: gb\_pl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	length DB ID	Description
1	5311	100.0	3380 6	CQ776470 Sequence
2	5311	100.0	3380 6	CS108225 Sequence
3	5311	100.0	3380 6	CS113027 Sequence

4	5311	100.0	3380 8	AF095844	AF095844 Homo sapi
5	5307	99.9	3434 6	CS033863	CS033863 Sequence
6	5307	99.9	3434 6	CS036676	CS036676 Sequence
7	5307	99.9	3434 6	CS042815	CS042815 Sequence
8	5307	99.9	3434 6	CS045628	CS045628 Sequence
9	5285	99.5	3372 6	AX300832	AX300832 Sequence
10	5285	99.5	3373 8	AY017378	AY017378 Homo sapi
11	4199	79.1	3078 9	AY075132	AY075132 Mus muscu
12	4194	79.0	3771 6	CQ777379	CQ777379 Sequence
13	4194	79.0	3771 9	AF374384	AF374384 Mus muscu
14	4080.5	76.8	3635 9	BC080200	BC080200 Mus muscu
15	3954	74.4	2540 8	BC078180	BC078180 Homo sapi
16	3358	63.2	2028 6	CQ729104	CQ729104 Sequence
17	2874	54.1	1776 8	AK056293	AK056293 Homo sapi
18	2874	54.1	1776 8	AK056293	AK056293 Homo sapi
19	2690.5	50.7	2468 9	BC025508	BC025508 Mus muscu
20	2442	46.0	1443 6	AX300841	AX300841 Sequence
21	2271.5	42.8	2050 9	BC004031	BC004031 Mus muscu
22	2072	39.0	1392 6	AR379503	AR379503 Sequence
23	2021	38.1	1204 8	BC007966	BC007966 Homo sapi
24	1995	37.6	1284 6	AX300838	AX300838 Sequence
25	1508	28.4	3458 5	BC073528	BC073528 Xenopus l
26	1357.5	25.6	2493 4	BT020952	BT020952 Bos tauru
27	1356	25.5	2326 9	AF316999	AF316999 Mus muscu
28	1356	25.5	2336 9	BC029209	BC029209 Mus muscu
29	1352	25.5	2584 8	BC014949	BC014949 Homo sapi
30	1352	25.5	2590 6	AR578595	AR578595 Sequence
31	1344	25.3	2613 6	BD157193	BD157193 Primer fo
32	1344	25.3	2613 6	AX878393	AX878393 Sequence
33	1344	25.3	2613 8	AK021416	AK021416 Homo sapi
34	1201.5	22.6	1258 6	AX098232	AX098232 Sequence
35	1201.5	22.6	1270 6	AX098236	AX098236 Sequence
36	1129	21.3	2250 6	AX834948	AX834948 Sequence
37	1129	21.3	2250 8	AK097669	AK097669 Homo sapi
38	1091.5	20.6	4372 6	CQ871510	CQ871510 Sequence
39	1091.5	20.6	4372 6	CS033815	CS033815 Sequence
40	1091.5	20.6	4372 6	CS042767	CS042767 Sequence
41	1091.5	20.6	4372 6	CS130808	CS130808 Sequence
42	1090.5	20.5	3065 6	CQ776474	CQ776474 Sequence
43	1090.5	20.5	3065 6	CQ776595	CQ776595 Sequence
44	1090.5	20.5	3065 6	CS108058	CS108058 Sequence
45	1090.5	20.5	3065 6	CS113026	CS113026 Sequence

ALIGNMENTS

RESULT 1  
CQ776470  
LOCUS CQ776470 3380 bp DNA linear PAT 11-MAR-2004  
DEFINITION Sequence 156 from Patent EP1394274.  
ACCESSION CQ776470  
VERSION CQ776470.1 GI:45379860  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1  
AUTHORS Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuhara,K.  
TITLE Methods of testing for bronchial asthma or chronic obstructive  
pulmonary disease  
JOURNAL Patent: EP 1394274-A 156 03-MAR-2004;  
Genox Research, Inc. (JP)  
FEATURES  
source location/Qualifiers  
1..3380  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores: 1.77e-312 Length: 3380  
Pred. No.: 1

Score:	5311.00	Matches:	1025
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-09-515-363C-2 (1-1025) X CQ776470 (1-3380)

QY	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	169	ATGTCGAATGGGTATTCACAGACGAAATTTCCGCTATCTCATCTCGTGCTTCAGGGCC	228
QY	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	229	AGGGTGAATAATGTACATCCACAGGTGAGCCTGTGTGCTGACTACCTGACCTTTCCTGCA	288
QY	41	GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	289	GAGGTGAAGAGCAGCATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGCAGCTTGA	348
QY	61	LeuLeuLeuSerThrLeuGlnLysGlyValTyrPheIleuGlyTyrThrArgGluPheVal	80
Db	349	CTGCTGCTGAGCACCTTGAGAGAAGGAGTCTGGCACCTTGGTTGGAATCGGAAATTCGTG	408
QY	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
Db	409	GAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCCGCCCTCATGAACCTGAGCTCAGC	468
QY	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	469	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATATCTCCAATGCTGAACCTC	528
QY	121	IeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
Db	529	CTTCAGCCCACTCTGGTGCACAAGCTTCTAGTAGAGAGCTCTTGATTAAGTGACATGAG	588
QY	141	GluGlnLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	589	GAGGAATGTTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAATGGAAT	648
QY	161	GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTyrPheSerAla	180
Db	649	GAATCAGGTGTAAGAGAGCTACTAAAAAGGATTGTGCAGAAAAAAGCTGTTCTTCGCA	708
QY	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	709	TTTCTGAATGTTCTTCGTCAAACAGGAAACAATGAATGTTGCCAAGATTAAACAGGCTCT	768
QY	201	AspCysSerGluSerAsnAlaGlnIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	769	GATTGCTCAGAAAGCAATGCAGAGATTGAGAATTTATCACAAGTTGATGTCCTCAAGTG	828
QY	221	GluGlnGlnLeuLeuSerThrThrValGlnProAsnLeuGlnLysGluValTyrGlyMet	240
Db	829	GAAAGCAACTTCTTTCAACCAAGTTCAGCCAAATCTGGAGAGAAGAGTCTGGGGCATG	888
QY	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValIleSerGluSerAspThr	260
Db	889	GAGAATACTCATCAGAAATCATCTTTTGACAGATTCTTCTGTAGTTTCAGAAATCAGACACA	948
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
Db	949	AGTTTGACAGAGAAAGTGTCACTGCTTAGATGAAGTCTTGACATTAACAGCAACATG	1008
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	1009	GGCAGTGAATTCAGCACCATGGGAAGTGAATTCAGATGAAGAAGATGTGCAGCAAGAGCA	1068
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaIleu	320
Db	1069	TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAAAATGGAAGTTGCCACGACCCCTTG	1128
QY	321	GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340

Db	1129	GAAGGAAGAAATATCATCATCTGCTCCCTACAGGAGTGAAAAACCAAGTGGCTGTT	1188
QY	341	TyrIleAlaLysAspHisIleuAspIleValylValAlaSerGluProGlyLysValIle	360
Db	1189	TACATTGCCAAGATCACTTAGACACAGAGAAAAAGCATCTGAGCCTGAAAAATTATA	1248
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380
Db	1249	GTTCTTGTCAAATAAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAACTATT	1308
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1309	TTGAAGAAATGGTATCGTGTATTGGATTAACTGGTGATACCCTGAATAATCATTTT	1368
QY	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1369	CCAGAGTTGTCAAGTCCCTGTGATATTATTCAGTACAGCTCAATCCTTGAAAACTCC	1428
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1429	CTCTTAACTTGGAAAAATGAGAGATGCTGGTGTTCATTGTGCAGACTTTCCCTCATTT	1488
QY	441	IleIleAspGluCysHisIleThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1489	ATCATTTGATGATGTCTATCACACCACAAAGACAGTGTATTAATACATCATGAGCAT	1548
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
Db	1549	TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCATGATTC	1608
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAla	500
Db	1609	CTTCCTCAGATACTGGGACTTAACAGCTTCACCTGCTGTGGAGGGGCCACGAGCAAGCC	1668
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1669	AAAGCTGAAGACACATTTTAAAACTAATGTGCCAATCTTGATGCATTTACTATTAAACT	1728
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	1729	GTTAAAGAAAACTTGATCAACTGAAAAACCAATAACAGAGCCATGCAAGAGTTGCC	1788
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
Db	1789	ATTGCAGATGCACACAGAGATCCATTTAAAGAGAACTTCTAGAAATATGACAAAG	1848
QY	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	1849	ATTCAACTTATTTGTCAAAATGAGTCCAAATGTGAGATTTTGGAACTCAACCTTATGAACA	1908
QY	581	TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1909	TGGGCCATTCAATGGAATAAAAAAGCTGCAAAAAAGAAATCGCAAGAACGTGTTGT	1968
QY	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	1969	GCAGAACATTTGAGGAAAGTACAATGAGGCCCTCAAAATTAATGACACAAATTCGAATGATA	2028
QY	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal	640
Db	2029	GATGCGTATACTCATCTTGAAACTTCTATTAATGAAGAAAGATAAGAGTTTGCAGTTC	2088
QY	641	IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	2089	ATAGAAGATGATAGTATGAGGGTGTGATGATGATGATGATGATGATGATGATGATGAG	2148
QY	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	2149	GATGATTTAAAGAAACCTTTGAAACTGATGAACACAGATAGATTCTCATGACTTTATTT	2208
QY	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700

Db 2209 TTTGAAAACAATAAATGTTGAAAAGGCTGGCTGAAAAACCGAATATGAAAAAG 2268  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720  
Db 2269 CTGACCAAAATTAGAAATACCAATAGAGCAATATAGAGACTGAGGAATCAGCACA 2328  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTripleThrGlu 740  
Db 2329 GGAATATCTTTACAAAAACAGCAGAGTGCAATATGCCGTTTCCAGTGGAATTAAGAA 2388  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisIleLeuIleGlyAlaGlyHisSer 760  
Db 2389 AATGAAAAATTGCTGAAGTAGAGTCAAAAGCCCACTCTGATTTGAGCTGCACACAGC 2448  
Qy 761 SerGluPheLysPrometThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2449 AGTGAGTTCAAAACCATGACACAGATGAACAAAAAGAGTCATTAGTAAATTTCCACT 2508  
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2509 GGAATAATCAATCTGCTTATCGCTACCAAGTGAGAGAAAGTCTGATATTTAAAGAA 2568  
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2569 TGTAACATGTTGATCCGTTATGCTCTCGTCAACCAATGAATAAGCATGTCAGGCCCT 2628  
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2629 GGTCCAGCCACAGACTGATGAGAGCACTACGTCCTGGTTGCTCACAGTGTTTCAGAGTT 2688  
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860  
Db 2689 ATCGAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGATAAGCTATACATTGT 2748  
Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2749 GTTCAAAATATGAAACCAAGAGATAGCTCATAGATTGTGAATTACAGATGCCAAAGT 2808  
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900  
Db 2809 ATAATGAAAAAGAAATGAAAAACCAAGAAATATTGCCAAACATTAACAAGATAACCCA 2868  
Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920  
Db 2869 TCACTAATACTTCTTCTTGCAAAACTGCAGTGCTGCTAGCCTGTTCTGGGGAAGATAATC 2928  
Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
Db 2929 CATGTAATTGAGAAAATGATCATGCTCAATATGACCCCAAGATTCAAAGAACTTTAATTC 2988  
Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 2989 GTAAGAGAAAAACAAGCACTGCAAAAGAAAGTGCAGACTATCAATAATAATGTCAAATC 3048  
Qy 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3049 ATCTGCAATGTGGCCAGGCTTGCGGAACAATGATGTCACAAAGCCTTAGATTGCCCT 3108  
Qy 981 CysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnTyr 1000  
Db 3109 TGTCTCAAAATAAGGAATTTTGTAGTGTTTCAAAAAATAATTCAACAAAGAAACAATAC 3168  
Qy 1001 LysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020  
Db 3169 AAAAAAGTGGTAGAATTACCTATCAATTTCCCAATCTTGACTAFTCAGAAATGCTGTTTA 3228  
Qy 1021 PheSerAspGluAsp 1025  
Db 3229 TTTAGTGATGAGGAT 3243  
RESULT 2  
CS108225  
LOCUS CS108225 3380 bp DNA linear PAT 22-JUN-2005

DEFINITION Sequence 233 from Patent WO2005051988.  
ACCESSION CS108225  
VERSION CS108225.1 GI:68147706  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
REFERENCE 1  
AUTHORS Abbas,A., Bodary,S., Clark,H., Schoenfeld,J., Williams,P.M.,  
Wood,W.I. and Wu,T.D.  
TITLE Compositions and methods for the treatment of systemic lupus  
erythematosus  
JOURNAL Patent: WO 2005051988-A 233 09-JUN-2005;  
Genentech, Inc. (US)  
FEATURES  
source 1..3380  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.77e-312 length: 3380  
Score: 5311.00 Matches: 1025  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 6 Gaps: 0  
US-09-515-363C-2 (1-1025) x CS108225 (1-3380)  
Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
Db 169 ATGTGCAATGGGTATTCACAGACAGAAATTTCCGCTATCTCATCTCGTGCTCAGGGCC 228  
Qy 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
Db 229 AGGGTGAATAATGTACATCCAGGTGAGCGCTGTGCTGACTAACCCTTCTGCTGCA 288  
Qy 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 289 GAGGTGAAGAGACAGATTGAGAGGACAGTCCGACCTCCGGAAACATGCAGGCAATTGAA 348  
Qy 61 LeuLeuSerThrLeuGluLysGlyValTyrHisLeuGlyTyrThrArgGluPheVal 80  
Db 349 CTGCTGTGAGCACCTTGAGAAAGGAGTCTGGCACTTGTTGACTCGGGAATTCGTG 408  
Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
Db 409 GAGGCCCTCCGAGAACCGGCAAGCCCTCTGGCCGCCCTACATGAACCTGAGCTCACG 468  
Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATAATCTCCAACCTGCTGAACCTC 528  
Qy 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140  
Db 529 CTTGAGCCCACTCTGTGTGACAAAGCTTCTAGTTAGACGCTTGATTAAGTCATGAGAG 588  
Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaAlaGluAsnAsnGlyAsn 160  
Db 589 GAGGAAGTGTGACAAATTGAAGACAGAAACCGGATGCTGCTGCAGAAAAACAATGAAAT 648  
Qy 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTyrPheSerAla 180  
Db 649 GAATCAGGTGTAAGAGAGCTATAAAAGGATGTGTGAGAAAGAAACTGTTCTCTGCA 708  
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 709 TTTCTGAATGTTCTTCTGTCAAACAGAAACAATGAATGAACTGTCCAAGAGTTACAGGCTCT 768

QY	201	ASPCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	769	GATTGCTCAGAAAGCAATGCAGAGATTGAGAATTATACAAAGTTGATGCTCAAGTG	828
QY	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240
Db	829	GAAAGCAACTCTTCAACACACAGTTCAGCCAAATCTGAGAGAGAGGTCTGGGCATG	888
QY	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValSerGluSerAspThr	260
Db	889	GAGATAACTCATCAGAAATCATCTTTTGACAGATTCTTCTGTAGTTTCAGAAATCAGACACA	948
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
Db	949	AGTTTGGCAGAGAGAGTGTCACTGCTTAGATGAAGTCTTGACATTAACAGACAATG	1008
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	1009	GGCAGTGAATTCAGGCACCATGGGAAGTGATTCAGATGAAGAATGTGGCAGCAGAGACA	1068
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
Db	1069	TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAAGTTGCCACGCCCTTG	1128
QY	321	GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
Db	1129	GAAAGGAAGATATCATCATCTGCTCCCTCACAGGAGTGAAAAACAGAGTGGCTGTT	1188
QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle	360
Db	1189	TACATTGCCAAGATCATCTTAGACAAGAAAAAGCATCTGAGCCTGGAAGAGTTATA	1248
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380
Db	1249	GTTCTTGTCAATAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCACCACTTT	1308
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1309	TTGAAGAAATGTTATCGTTATTGATTAGTGGTGATACCCCAACTGAAATATCATTT	1368
QY	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1369	CCAGAAGTTGTCAAGTCCCTGTGATATTATTATCAGTACAGCTCAAACTCTGAAAACTCC	1428
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1429	CTCTTAACTTGAAAAATGAGAGAAGATGCTGTTGTTCAATTGTCAGACTTTTCCCTCATT	1488
QY	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1489	ATCATTGATGAATGTATCATCACCAACAAGAGCAGTGTATTAATTAACATCATGAGCAT	1548
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
Db	1549	TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCACTGATTTCCC	1608
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500
Db	1609	CTTCCTCAGATACTGGGACTTAACAGCTTCACCTGTTGTTGAGAGGGCCACGAGCAAGCC	1668
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1669	AAAGCTGAAGAACACATTTTAAACTATATGTGCCAATCTTGATGCATTTACTATTAAACT	1728
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	1729	GTTAAAGAAAACTTGATCACTGAAAAAACCAATACAGAGCCATGCAAGAAAGTTGCC	1788
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
Db	1789	ATTGCAGATGCAACCAAGAGAGATCCATTTTAAAGAGAACTTCTAGAATATATGACAAG	1848
QY	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580

[illegible]

Db 2929 CATGTAATTGAGAAAATGCATCAGTCAATATGACCCGAGAATTCAAGGAACTTTACATT 2988  
Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 2989 GTAAGAGAAAAACAAGCACCTGCAGAAAGAGTGTGCCGACTATCAATTAATGATGAAATC 3048  
Qy 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3049 ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAGGCTTAGATTGCCCT 3108  
Qy 981 CysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnTyr 1000  
Db 3109 TGTCTCAAAATAAGGAATTTGTAGTGCTTTCAAAAATAATTCAACAAAGAAACAATAC 3168  
Qy 1001 LysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGlyCysLeu 1020  
Db 3169 AAAAAGTGGGTAGAAATTACCTATCACTTCCCAATCTTGACTATTGAGAATGCTGTTTA 3228  
Qy 1021 PheSerAspGluAsp 1025  
Db 3229 TTTAGTGATGAGGAT 3243

RESULT 3  
CS113027 3380 bp DNA linear PAT 24-JUN-2005  
LOCUS CS113027  
DEFINITION Sequence 45 from Patent WO2005054507.  
ACCESSION CS113027  
VERSION CS113027.1 GI:68224599  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE 1  
AUTHORS Corfe, B. and Chirakkal, H.  
TITLE Gene screen  
JOURNAL Patent: WO 2005054507-A 45 16-JUN-2005;  
University of Sheffield (GB)  
FEATURES  
source location/Qualifiers  
1..3380  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:  
Pred. No.: 1.77e-312 length: 3380  
Score: 5311.00 Matches: 1025  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 6 Gaps: 0

US-09-515-363C-2 (1-1025) x CS113027 (1-3380)

Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
Db 169 ATGTGCAATGGGTATTCACAGACGAGAATTTCCGCTATCTCATCTCGTTCAGGGCC 228  
Qy 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
Db 229 AGGGTGAATAATGTACATCCAGGTGAGCCTGTGCTGACTAACCCTTCTGCTTGCCTGCA 288  
Qy 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 289 GAGGTGAAGAGCAGATTCAAGAGACAGTGCACACTCCGGGAACATGACAGCAGTTGAA 348  
Qy 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrPheIleGlyTyrThrArgGluPheVal 80  
Db 349 CTGCTGCTGAGCACCTTGAGAGGAGAGTCTGGCACCTTGAGTGGGAAATTCGTG 408  
Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100

Db 409 GAGGCCCTCCGAGAACCGGAGCCCTCTGCGCCGCTACATGAACCTGTGACTCACG 468  
Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACCTC 528  
Qy 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140  
Db 529 CTTGACCCCACTCTGGTGCACAGCTTCTAGTTAGACAGCTCTTGATTAAGTGCATGGAG 588  
Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaValGluAsnAsnGlyAsn 160  
Db 589 GAGGAATGTTGACAATTGACAGACAGAAAACCGAATTGCTGCTGCAGAAAACAATGGAAT 648  
Qy 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTyrPheSerAla 180  
Db 649 GAATCAGGTGTAGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAACGTGTTCTCTGCA 708  
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 709 TTTCTGAATGTTCTTCGTCAACAGGAACAATGAATGTGTCCAAGAGTTTACAGGCTCT 768  
Qy 201 AspCysSerGluSerAsnAlaGlnIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 769 GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTATCACAAAGTTGATGCTCAAGTG 828  
Qy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTyrGlyMet 240  
Db 829 GAAGCAACTTCTTTCAACCAACAGTTCAGCCAAATCTGAGAAAGAGGTCTGGGGCATG 888  
Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValAspGlySerAspThr 260  
Db 889 GAGATAACTCATCAGAACTCATCTTTTGACAGATTCTTCTGAGTTTCAAGATCAGACACA 948  
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 949 AGTTTGCAGAGAAGAGAGTGTACGCTGTTAGATGAAGTCTTGACATTAACGACAATG 1008  
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
Db 1009 GGCAGTGATTCAAGCACCATGGGAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGACA 1068  
Qy 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
Db 1069 TCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTG 1128  
Qy 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340  
Db 1129 GAAGGGAAGATATCATCATCTGCCCTCCCTACAGGAGTGGAACCAAGAGTGGCTGTT 1188  
Qy 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360  
Db 1189 TACATTGCCAAGGATCATTAGACAAGAAAGAAAAGCATCTGAGCCTGGAAGTATATA 1248  
Qy 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380  
Db 1249 GTTCTTGTCATTAAGGTACTGCTAGTTGAACAGCTCTTCCGAAAGAGTTCCAACCATTT 1308  
Qy 381 LeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400  
Db 1309 TTGAAGAAATGTATCGTTATTGATTAAAGTGTGATACCCAACTGAATAATATCATTT 1368  
Qy 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
Db 1369 CCAGAAAGTTGTCAAGTCTGTGATATTATATCAGTACAGCTCAAAATCCTGAAAACTCC 1428  
Qy 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
Db 1429 CTCTTAACCTTGAAAATGGAAGAAGATGCTGTTCAATTGTCAAGACTTTCCCTCATTT 1488  
Qy 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460

Db 1489 ATCATTTGATGATGTGCATCACACCAACAAAGACGAGTGTATATACATCATGAGGCAT 1548  
Qy 461 TTYLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480  
Db 1549 TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAGTGAATTCCTC 1608  
Qy 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500  
Db 1609 CTTCCTCAGATACTGGAGACTAACGCTTCACCTGGTGTGGAGGGGCCACGAAACCAAGCC 1668  
Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
Db 1669 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACT 1728  
Qy 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540  
Db 1729 GTTTAAAGAAAACTTGATCACTGAAAAACCAATACAGAGCCATGCAGAAAGTTTGGCC 1788  
Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560  
Db 1789 ATTGCAGATGCACACGAGAAAGATCCATTTAAAGAGAACTTCTAGAAATATGACAAAG 1848  
Qy 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580  
Db 1849 ATTCAAACTTATTGTCAAAATGATGCCAATGTCAGATTTTGAACCTCAACCTTATGAACAA 1908  
Qy 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
Db 1909 TGGGCCATTCAAATGGAATAAAAGCTGCATAAAAGAAATCGCAAAAGAACGTGTTCGT 1968  
Qy 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
Db 1969 GCAGAACATTTGAGGAGTACATGAGGCCCTACAAATTATGACACACAAATTCGAATGATA 2028  
Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640  
Db 2029 GATGCCGTAATCTCATCTTGAACTTTCTATAATGAAAGAGAAAGATAAGAGTTTGCACTC 2088  
Qy 641 IleGluAspAspSerAspGluGlyLysAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 2089 ATAGAGATGATAGTATGATGAGGGTGTGATGATGAGTATTGTGATGTGATGAGATGAG 2148  
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2149 GATGATTTAAAGAAACCTTTGAACTGGATGAAACAGATAGATTCTCATGACTTTATT 2208  
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2209 TTTGAAAAACATAAATGTGTAAGAGGCTGGCTGAAAAACCAAGATATGAAAAATGAAGAAG 2268  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluLysSerAlaArg 720  
Db 2269 CTGACCAAATTTAAGAAATACCATATGAGCAATATATAGAACTGAGGAATCAGCAGCA 2328  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740  
Db 2329 GGAATGATCTTTACAAAAACAGCAGAGTGCATATGCGCTTTCCAGTGGATTACTGAA 2388  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisIleLeuIleGlyAlaGlyHisSer 760  
Db 2389 AATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGGAGCTGCACAGC 2448  
Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2449 AGTGAATTCAAAACCCATGACACAGATGAACAAAAAGAAATCATTAAGTAATTTGCACT 2508  
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2509 GGAATAATCAATCTGCTTATCGCTACCAACAGTGGCAGAAAGAGTCTGGATTATTAAGAA 2568  
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2569 TGTAACATTTGTTATCGTTATGGTCTGCTGACCAATGAAATAGCCATGGTCCAGGCCCT 2628

Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2629 GGTGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTGTCTCACAGTGGTTCAAGAGTT 2688  
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860  
Db 2689 ATCGAACATGAGACAGTTAATGATTTCCGAGAGAGATGATGATATAAGCTATACATTGT 2748  
Qy 861 ValGluAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2749 GTTCAAAATATGAACCAAGAGAGTATGCTCATAGATTTTGAATTACAGATGCAAGT 2808  
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900  
Db 2809 ATATGCAAAAAGAAATGMAAACCAAGAGAAATATTGCCAAGCATTTACAGAAATAACCA 2868  
Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValIleAlaCysSerGlyGluAspIle 920  
Db 2869 TCACCTAATACTTCTTTCGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGAGATATC 2928  
Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
Db 2929 CATGTAAATTGAGAAATGCAATCACTCAATATGACCCAGAAATTCAGAAACTTTACATT 2988  
Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 2989 GTAAGCAAAAACAAGACACTGCATAAAGAGTGTGCCAGCTATCAATAATAATGTGAATC 3048  
Qy 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3049 ATCTGCATAATGTGGCCAGGCTTTGGGGAACAATGATGTGCACAAAGCCTTAGATTGCC 3108  
Qy 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000  
Db 3109 TGTCTCAAAATTAAGAAATTTGTAGTGTGTTTCAAAAATAATTCAACAAAGAAACAATAC 3168  
Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGlyCysLeu 1020  
Db 3169 AAAAAGTGGTAGAATTAACCTATCACATTTCCCAATCTTGACTATTCAAGATGCTGTTTA 3228  
Qy 1021 PheSerAspGluAsp 1025  
Db 3229 TTTAGTGATGAGGAT 3243  
RESULT 4  
AF095844 3380 bp mRNA linear PRI 23-JAN-2002  
LOCUS  
DEFINITION Homo sapiens melanoma differentiation associated protein-5 (MDA5)  
ACCESSION AF095844  
VERSION AF095844.1 GI:11344593  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 3380)  
Kang,D.C., Gopalkrishnan,R.V., Wu,Q., Jankowsky,E., Pyle,A.M. and  
Fisher,P.B.  
TITLE mda-5: An interferon-inducible putative RNA helicase with  
double-stranded RNA-dependent ATPase activity and melanoma  
growth-suppressive properties  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (2), 637-642 (2002)  
PUBMED 11805321  
REFERENCE  
AUTHORS 2 (bases 1 to 3380)  
Kang,D.-C. and Fisher,P.B.  
TITLE Direct Submission  
JOURNAL Submitted (30-SEP-1998) Urology, Columbia University, PES, 630  
W168th, New York, NY 10032, USA  
FEATURES  
source 1..3380

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HO-1"
/cissue_type="melanoma"
1..3380
/gene="MDA5"
169..3246
/gene="MDA5"
/codon_start=1
/product="melanoma differentiation associated protein-5"
/protein_id="AAG34368.1"
/db_xref="GI:11344594"
/translation="MSGYSTDENFRYLISCFRAVKMYIQVEPVLDTLPFAEYKE
QIQRTVATSGNMQAVEILLSTLEKGVHLGWTREFEALRRTGSPLAARYMPELTDL
PSPSFENAHDEYLOLNLLOPTLVLDKLCMBEELTIEDNRRIAAENNGN
ESGVRELLKRIVQKENWFSAPLNLVRLQGTNNELVQELTGSDCSESNAELENLSQVDP
QVEEQLSTTVQPNLEKEVWMENNSSESFADSSVSESOTSIAEGSVSCLDESLGH
NSNMGSDSGTMGSDSDSENVAAASPEPELQLRPYQMEVAQPALEGKNIICLPYSG
KTRVAVYIAKDLDKKKASEPGKIVLVNKKVLVEQLFRKEPQPLKKMYRVLGLSG
DTQKISFPEVVKSCDIIISTAQILNSLNLNENGADAGVOLSPSLIIDECHHTNK
EAVYNNIMRHYLMQKLKNRLKKNKPVIPLOLGLTASPGVGATKQAKAEHLIK
LCANLDAFTIKVKENLDQKNQIOEPCKFAIADATREDPFKEKLEIMTRIQTCCO
MSPMSDFGTQPYEQWAIQMEKKAQKGRKERVCAEHLRKYNEALQINDTIRMIDAYT
HLETFYNEEKOKKFAVIEDSDDEGDEYCDGEDDEDLKKPLKLDETDRFLMTLFE
NNKMLKRLAENPEYENEKLTKLRLTIMEQYTRTEESARGIIFTKTRQSAVALSQWITE
NEKFAEVGKAAHHLIGAGHSSEPKPMTQNEQKEVISKFTGKNIILATVAEEGLDI
KECNIVIRYGLVTNEIAMVQARGRADESTVYLVAHSGSVIEHETVNDFREKMYK
AIHCVQNMKPEEYAHKILBLQMSIMEKKMKTARNIAKHYNKPSLITFLCKNCVLA
CSGEDIHVIEKMHVNMTPFEKELYIRENKALQKCADYQNGEIIICKGOAGTMM
VHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVELPTFPNLDSSECCLFSDSD"
```

ORIGIN

Alignment Scores:

Pred. No.:	1.77e-312	Length:	3380
Score:	5311.00	Matches:	1025
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-09-515-363C-2 (1-1025) x AF095844 (1-3380)

Qy	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	169	ATGTGCAATGGTATTCCACAGACGAAATTCGCTACTCATCTCGTGTTCAGGGCC	228
Qy	21	ArgValIysMetTyrIleGlnValGluProValIleuAspTyrLeuThrPheIleuProAla	40
Db	229	AGGTCGAAATGTACATCCAGGTGGAGCCTGTGCTGACTGACCTTCTGCCTGCA	288
Qy	41	GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	289	GAGGTGAAGAGCAGATTTCAGAGCACAGTCGCCACCTCCGGGAACATGAGGCAGTTGAA	348
Qy	61	LeuLeuLeuSerThrLeuGluIysGlyValTyrPheIleuGlyTyrThrArgGluPheVal	80
Db	349	CTGCTGCTGAGCACCTTGAGAGGGAAGTCTGGCACCTTGCTTGACTCGGGAATTCTGTG	408
Qy	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
Db	409	GAGGCCTCCGAGAACCGGAGCCCTCTGCGCCGCCCTACATGAACCTGAGCTCACG	468
Qy	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	469	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACCTC	528
Qy	121	LeuGlnProThrLeuValAspIysLeuLeuValArgAspValLeuAspIysCysMetGlu	140
Db	529	CTTCAGCCCACTCTGGTGACAAGCTTCTAGTTAGACGCTCTTGAGATAAGTCATGGAG	588
Qy	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160

Db	589	GAGGAACCTGTTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAACAATGGAAT	648
Qy	161	GluSerGlyValArgGluLeuLeuIysArgIleValGlnIysGluAsnTrpPheSerAla	180
Db	649	GAATCAGGTGTAAAGAGACTAATAAAAGATTGTGCAGAAAGAAAACCTGTTCTCTGCA	708
Qy	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	709	TTTCTGAATGTTCTTCGTCAAAACAGAAACAAATGAACCTTGTCCAAGAGTTAAACAGCTCT	768
Qy	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	769	GATTGCTCAGAAAGCAATGCAGAGATTGAGATTATCACAGTTGATGGTCTCAAGTG	828
Qy	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIysGluValTrpGlyMet	240
Db	829	GAAAGCAACTCTTTTCAACCAAGTTCAGCCAAATCTGGAGAAAGAGGCTCTGGGGCATG	888
Qy	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValIleSerGluSerAspThr	260
Db	889	GAGATAAATCATCAGAAATCATCTTTTGAGATTCTTCTGTAGTTTCAGAAATCAGACACA	948
Qy	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerIleuGlyHisAsnSerAsnMet	280
Db	949	AGTTTGGCAGAGGAAGTCTCAGCTGCTTAGTGAAGTCTTGACATPAACAGCAACATG	1008
Qy	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	1009	GGCAGTGAATTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGATGTGGCAGCAAGACA	1068
Qy	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
Db	1069	TCCCGGAGCCAGAACTTCAGCTCAGGCCCTTACCAATGGAAAGTTGCCAGCCAGCCTTG	1128
Qy	321	GluGlyIysAsnIleIleIleCysLeuProThrGlySerGlyIysThrArgValAlaVal	340
Db	1129	GAAAGGAAGATATATCATCATCTGCCCTCCCTACAGGAGTGAAAAACCAAGATGGCTGTT	1188
Qy	341	TyrIleAlaIysAspHisLeuAspIysIysIysAlaSerGluProGlyIysValIle	360
Db	1189	TACATTGCCAAGATCATCTTAGACAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATA	1248
Qy	361	ValLeuValAsnIysValLeuLeuValGluGlnLeuPheArgIysGluPheGlnProPhe	380
Db	1249	GTTCTTGTCAATAAGTACTGCTAGTTGACACGCTCTCCGACAGAGTCCCAACCATTT	1308
Qy	381	LeuIysIysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuIysIleSerPhe	400
Db	1309	TTGAGAAATGTTATCGTTATTGATTAGTGTGATACCCCAACTGAAAAATATCATTTT	1368
Qy	401	ProGluValValIysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1369	CCAGAAAGTTGCAAGTCTCTGTATATTATATATATATATATATATATATATATATATAT	1428
Qy	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1429	CTCTTAAACTTGAAAAATGGAAGAGATGCTGTGTTCAATTGTACAGACTTTTCCCTCATT	1488
Qy	441	IleIleAspGluCysHisHisThrAsnIysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1489	ATCATTTGATGAATGTCTACACCAACAAGACAGAGTATATAATACATCATGAGGCAT	1548
Qy	461	TyrLeuMetGlnIysLeuIysAsnAsnArgLeuIysGluAsnIysProValIlePro	480
Db	1549	TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAAACCAACCAAGTATTCCC	1608
Qy	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrIysGlnAla	500
Db	1609	CTTCCTCAGATACTGGACTAAACAGCTTCACTGTGTGTGAGGGGCCACGAAAGCAAGCC	1668
Qy	501	IysAlaGluGluHisIleLeuIysLeuCysAlaAsnLeuAspAlaPheThrIleIysThr	520
Db	1669	AAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATGATTAATTAATAAACT	1728

QY	521	VallYsGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	1729	GTTAAAGAAAACCTTGATCAACTGAAAAAACCAATACAGAGCCATGCAAGAGTTTGCC	1788
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
Db	1789	ATTGACAGATGCAACCGAGAAAGATCCATTTAAGAGAACTTCTAGAAATTAATGACAAAG	1848
QY	561	IleGlnThrTrpCysGlnMetSerPrometSerAspPheGlyThrGlnProCysGlnGln	580
Db	1849	ATTCAACTTATGTGCAATAGAGTCCCAATGTCAAGTTTGGAACTCAACCCCTATGAACAA	1908
QY	581	TrpAlaIleGlnMetCylulysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1909	TGGGCCATTCAATGGAAAAAAAAGCTGCAAAAAAGAAATCGCAAGAAGCGTTTGT	1968
QY	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	1969	GCAGAACATTTGAGGAGTACCAATGAGGCCCTCAAAATTAAATGACACAAATTCGAATGATA	2028
QY	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGlnGluLysAspLysLysPheAlaVal	640
Db	2029	GATGCGTACTCATCTGAACTTCTATATAGAGAAAGATAAGAACTTTCACAGTC	2088
QY	641	IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	2089	ATAGAAGATGATAGTGATGAGCGGTGATGATGATGATGATGATGATGATGATGATGAG	2148
QY	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	2149	GATGATTTAAAGAAACCTTTGAAACTGATGAACAGATAGATTCTCATGACTTTATT	2208
QY	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db	2209	TTTGAAAAACAATAAATGTGAAAAAGGCTGGCTGAAAAACCAAGATATGAAAAAG	2268
QY	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluLysSerAlaArg	720
Db	2269	CTGACCAATTAAGAAATACCATATATGAGCAATATATCTAGAGACTGAGGAATCAGCAGA	2328
QY	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu	740
Db	2329	GGAATAATCTTTACAAAAACAGACAGAGTGCATATGCCCTTCCACAGTGATTACTGAA	2388
QY	741	AsnGluLysPheAlaGluValGlyValLysAlaHisLysLeuIleGlyAlaGlyHisSer	760
Db	2389	AATGAAAAATTGCTGAAGTAGAGTCAAAAGCCCACTGTGATTGGAGCTGACACAGC	2448
QY	761	SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr	780
Db	2449	AGTGAGTTCAAACCCCATGACACAGAATGAACAAAAAGAGTCATTAGTAATTTCCACT	2508
QY	781	GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu	800
Db	2509	GGAATAATCAATCTGCTTATCCGCTACACAGTGCAGAGAGAGTCTGGATATTAAAGAA	2568
QY	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	820
Db	2569	TGTAAACATTGTATCCGTTATGCTCTGTCACCAATGAATAAGCCATGTGCCAGGCCGT	2628
QY	821	GlyArgAlaArgAlaAspGlnSerThrTyrValLeuValAlaHisSerGlySerGlyVal	840
Db	2629	GGTGAGCCAGAGCTGATGAGACACCTACGTCCTGGTGTCTCACAGTGTTCAGGAGTT	2688
QY	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860
Db	2689	ATCGAACATGAGACAGTTATGATTTCCGAGAGAGATGATGATATAAGCTATACATTGT	2748
QY	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer	880
Db	2749	GTTCAAAATATAGAAACCGAGAGTATGCTCATAGATTTTTGGAAATTACAGATGCAAAAGT	2808

Qy	881	IlleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisIleTyrLysAsnAsnPro	900			
Db	2809	ATAATGGAAGAAAGAAATGAAAAACCAAGAGAAATATTGCCAAGCATTTACAAGATAACCCA	2868			
Qy	901	SerLeuIleThrPheLeuCySLysAsnCySerValLeuAlaCySerGlyGluAspIle	920			
Db	2869	TCACTAATAACTTTCCTTGCAAAAACCTGCAGTGTCTAGCCCTGTTCTGGGGAAGATATC	2928			
Qy	921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheGlyLeuTyrIle	940			
Db	2929	CATGTAATTGAGAAATGCATCACGTCATATGACCCAGAAATTCAGGAACCTTACATT	2988			
Qy	941	ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle	960			
Db	2989	GTAAGAGAAAACAAGACACTGCAAAAGAAAGTGTCCGACTATCAAAATTAATGTGAAATC	3048			
Qy	961	IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro	980			
Db	3049	ATCTGCAAAATGTGCCACAGGCTTGGGGAACAATGATGTCACAAAGCCTTAGATTGGCCT	3108			
Qy	981	CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr	1000			
Db	3109	TGTCTCAAAATAGAGAAATTTGTAGTGTGTTTCAAAAATAATTCAACAAGAAACAATAC	3168			
Qy	1001	LysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGlyCysLysLeu	1020			
Db	3169	AAAAAGTGGTAGAATTACCTATCACATTCCCAATCTTGACTATTCAAGATGCTGTTTA	3228			
Qy	1021	PheSerAspGluAsp	1025			
Db	3229	TTTAGTGATGAGGAT	3243			
RESULT 5						
CS033863		3434 bp	DNA			
LOCUS	CS033863		linear			
DEFINITION	Sequence 3369 from Patent WO2005016962.		PAT 10-MAR-2005			
ACCESSION	CS033863					
VERSION	CS033863.1	GI:60732714				
KEYWORDS	.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;					
	Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Abbas,A., Clark,H., Ouyang,W., Williams,M.P., Wood,W.I. and Wu,T.D.					
TITLE	Compositions and methods for the treatment of immune related diseases					
JOURNAL	Patent: WO 2005016962-A 3369 24-FEB-2005;					
	Genentech, Inc. (US)					
FEATURES	Location/Qualifiers					
Source	1..3434					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
ORIGIN						
Alignment Scores:						
Pred. No.:	3,14e-312	Length:	3434			
Score:	5307.00	Matches:	1024			
Percent Similarity:	100.0%	Conservative:	1			
Best Local Similarity:	99.9%	Mismatches:	0			
Query Match:	99.9%	Indels:	0			
DB:	6	Gaps:	0			
US-09-515-363C-2 (1-1025) x CS033863 (1-3434)						
Qy	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20			
Db	223	ATGTGCAATGGGTATTCACACAGACGAGAAATTCGCTATCTCATCTCGTTCAGGGCC	282			
Qy	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40			

Db 283 AGGGTGAATAATGTACATCCAGGTGAGCCCTGTGCTGACTACCTGACCTTTCTGCTGCA 342  
Qy 41 GluVallySGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 343 GAGGTGAAGAGCAGATTCAAGAGCAGCTGCCACTCCGGAAACATGACAGCACTTGAA 402  
Qy 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrPheIleuGlyTyrThrArgGluPheVal 80  
Db 403 CTGCTGCTGAGCACCTTGAGAGAGGAGTCTGGCACCTTGCTTGACTCGGGAAATTCGTG 462  
Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
Db 463 GAGGCCCTCCGGAGAACCCGCCCTCTGGCCGCCCTACATGAACCTGAGCTCACG 522  
Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 523 GACTTGCCCTCTCCATCGTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 582  
Qy 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140  
Db 583 CTTGACCCCACTCTGTGAGCAAGCTTCTAGTAGACGTCCTTGATTAAGTCATGAGAG 642  
Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
Db 643 GAGGAATGTTGACAAATTGAGACAGAAACCGGATTGCTGCTGCAGAAACATGGAAT 702  
Qy 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTyrPheSerAla 180  
Db 703 GAATCAGGTGTAAAGAGAGCTACTAAAGAGATTGTGCAGAAAGAAACTGGTCTCTGCA 762  
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 763 TTTCTGAATGTTCTTCTGTAACAGAAACAATGAACTTGTCCAAAGAGTTAAcAGGCTCT 822  
Qy 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 823 GATTGCTCAGAAAGCAATGACAGATTGAGAAATTATCAcAAGTTGATGCTCTCAAGTG 882  
Qy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTyrGlyMet 240  
Db 883 GAAGAGCAACTTCTTCAACCAAGTTCAGCCAATCTGGAGAGAGGCTGCGGCATG 942  
Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
Db 943 GAGAAATTACTCATCAGAAATCATCTTTTGcAGATTCTTCTGTAGTTTCAGAATCAGACACA 1002  
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 1003 AGTTTGcAGAAAGAAAGTGTcAGCTGCTTAGATGAAGTCTTGACATTAACAGCAACATG 1062  
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
Db 1063 GGCAGTGAATTCAAGCACCATGGGAAGTATTcAGATGAAGAAATGTGGCAGCAAGACA 1122  
Qy 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
Db 1123 TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCAGCCAGCCTTG 1182  
Qy 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340  
Db 1183 GAAGGGAAGAAATATCATCATCTTGCTCCCTTACAGGGAGTGAAGAAAACcAGAGTGCCTGT 1242  
Qy 341 TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle 360  
Db 1243 TACATTGCCAAGGATCACTTAGACAGAAAGAAAGCATCTGAGCCTGGAAGATTATA 1302  
Qy 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380  
Db 1303 GTTCTGTCAATAAGTACTGTAGTTGAACAGCTCTCCGCAAGAGATTCCAACTATT 1362  
Qy 381 LeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400  
Db 1363 TTGAAGAAATGTATCGTGTATTGATTAAAGTGTGATACCcAACTGAAAAATATCATTT 1422

Qy 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
Db 1423 CCAGAAAGTTGTAAGTCTGTGATATATTATTCAGTACAGCTCAAACTCTTGAAAACTCC 1482  
Qy 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
Db 1483 CTCTTAAACTTGGAATAAGAGAGATGCTGGTGTTCATTTGTcAGACTTTTCCCTCATTT 1542  
Qy 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460  
Db 1543 ATCATTTGATGAATGTTCACACACCACAAAGAGAGTGTATATATACATCATGAGGCAT 1602  
Qy 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysProValIlePro 480  
Db 1603 TATTGATGCAGAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACcAGTATTCCC 1662  
Qy 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyValaThrLysGlnAla 500  
Db 1663 CTTGCTCAGATACTGGGACTAAcAGCTTCACTGGTGTGGAGGGGCCAGAAcCAAGCC 1722  
Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
Db 1723 AAAGCTGAAGACACATTTTAAACTATGTGCCAACTCTTGATGCATTTCTATTAAAACT 1782  
Qy 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540  
Db 1783 GTTAAAGAAAACTTGATCAACTGAAAAACCAATACAGAGCCATGCAGAAAGTTTGCC 1842  
Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560  
Db 1843 ATTGCAGATGCAAACcAGAGAGATCCATTTAAGAGAACTTCTAGAAATTAATGACAAAG 1902  
Qy 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580  
Db 1903 ATTCAACTTATTGTCAATGAAGTCCAAATGTcAGATTTTGGAACCTCAACCTTATGAACAA 1962  
Qy 581 TyrAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
Db 1963 TGGGCCATTCAATGGAAAAAAAGCTGCAAAAGAGAAATCGCAAGAGAGTGTGT 2022  
Qy 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
Db 2023 GCAGAACATTTGAGGAAGTACAATGAGGCCCTTACAAATTAATGACACAAATTCCGAATGATA 2082  
Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640  
Db 2083 GATGCGTATACTCATCTTGAACCTTTCTATTAAGAGAAAGATAAGAGTTTGcAGTc 2142  
Qy 641 IleGluAspAspSerAspGluGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 2143 ATAGAGATGATAGTAGTAGGGTGTGATGATGAGTATTGTGATGGTGAAGATGAG 2202  
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2203 GATGATTTAAAGAAACCTTTGAAACTGGATGAACAGATTGCTCATGACTTATATT 2262  
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2263 TTTGAAGAACATTAATATTGAAAAAGCTGGCTGAAAAACcCAGAAATATGAAGAAAG 2322  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720  
Db 2323 CTGACCAAAATTAAAGAAATCAATAATGAGCAATATATACTAGACTGAGGAATCAGACGA 2382  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGlu 740  
Db 2383 GGAATAATCTTTACAAAAACAGACAGAGTGCATATGCGCTTCCcAGTGAATTACTGAA 2442  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760  
Db 2443 AATGAAAAATTTGCTGAGATAGAGAGTCAAAAGCCCACTGTGATTGAGCTGGACACAGC 2502

Qy	761	SerGluPheIysProMetThrGlnAsnGluGlnIlySerGluValIleSerIysPheArgThr	780
Db	2503	AGTGAGTTCAAACCCATGACACAGAATGAACAAAAGAAGTCATTAGTAAATTTCCGACT	2562
Qy	781	GlyIlySILEAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleIysGlu	800
Db	2563	GGAAAAATAAATCTGCTTATCGCTACACAGTGGCAGAAAGGCTTGATATTAAAGAA	2622
Qy	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	820
Db	2623	TGTAACATTGTATTCCGTTATGCTCTCGTCACCAATGAATAAGCCATGCTCAGGCCCGT	2682
Qy	821	GlyArgAlaArgAlaAspGluSerThrTyrValIleuValAlaHisSerGlySerGlyVal	840
Db	2683	GGTCGAGCCAGAGCTGATGAGACACCTACGTCCTGGTGTCTCACAGTGGTTCAGAGATT	2742
Qy	841	IleGluHisGluThrValAsnAspPheArgGluIlyMetMetTyrIlyValIleHisCys	860
Db	2743	ATCGAACATGAGACAGTAAATGATTTCCGAGAGAAGATGATATAAGCTATACATTTGT	2802
Qy	861	ValGlnAsnMetIysProGluGluTyrAlaHisIlySILEuGluLeuGlnMetGlnSer	880
Db	2803	GTTCAAAATATGAAACCAAGAGAGTATGCTCATAGATTTTGGATTAACAGTCCAAAGT	2862
Qy	881	IleMetGluIlySlySmetIysThrIlyAsnIleAlaIlyshiSTyrIlyAsnAsnPro	900
Db	2863	ATTAATGAAAAAGAAATGAAAAACAAGAAATATTGCCAAGCATTACAAGATTAACCCA	2922
Qy	901	SerLeuIleThrPheLeuCysIlyAsnCysSerValIleuAlaCysSerGlyGluAspIle	920
Db	2923	TCACTAATAACTTTCCTTGGCAAAAACGTCAGTGTGTAGCCTGTTCTGGGGAAGATTATC	2982
Qy	921	HisValIleGluIlySmetHisIlyValAsnMetThrProGluPheIysGluLeuTyrIle	940
Db	2983	CATGTAATTGAGAAATGATCATCGTCAATATGACCCAGAATTCAAGGAATTACATT	3042
Qy	941	ValArgGluAsnIlySAlaLeuGlnIlySlySAlaAspTyrGlnIleAsnGlyGluIle	960
Db	3043	GTAAGAGAAAAACAAGCAGCTGCAAAAAGAAGTGTGCCGACTATCAATATAATGTTGAATC	3102
Qy	961	IleCysIlySlySylGlnAlaTyrGlyThrMetMetValHisIlySylLeuAspLeuPro	980
Db	3103	ATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGCCCT	3162
Qy	981	CysLeuIlySILEArgAsnPheValValPheIlyAsnAsnSerThrIlySlySglntYr	1000
Db	3163	TGTTCAAAATAAGGAATTTTGTAGTGTGTTTCAAAAATAATTCAACAAAGAAACAATAC	3222
Qy	1001	IlyIlySerTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu	1020
Db	3223	AAAAAGTGGGTAGAATTACCTATCATATTCCCAATCTTGACTATTCAGAATGCTGTTTA	3282
Qy	1021	PheSerAspGluAsp 1025	
Db	3283	TTTAGTGATGAGGAT 3297	
RESULT 6			
LOCUS	CS036676	3434 bp	DNA
DEFINITION	Sequence 6182 from Patent WO2005016962.	linear	PAT 10-MAR-2005
ACCESSION	CS036676		
VERSION	CS036676.1	GI:60734150	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1	Abbbs,A., Clark,H., Ouyang,W., Williams,M.P., Wood,W.I. and Wu,T.D.	
AUTHORS		Compositions and methods for the treatment of immune related	
TITLE		diseases	
JOURNAL		Patent: WO 2005016962-A 6182 24-FEB-2005;	

FEATURES			
source			
Genentech, Inc. (US)			
Location/Qualifiers			
1..3434			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Alignment Scores:			
Pred. No.:	3,14e-312	Length:	3434
Score:	5307.00	Matches:	1024
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	99.9%	Mismatches:	0
Query Match:	99.9%	Indels:	0
DB:	6	Gaps:	0
US-09-515-363C-2 (1-1025) x CS036676 (1-3434)			
Qy	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	223	ATGTGCAATGGGTATTCCACAGACGAGAATTCCGCTATCTCATCTCGTGTTCAGGGCC	282
Qy	21	ArgValIlySmetTyrIleGlnValGluProValIleuAspTyrLeuThrPheLeuProAla	40
Db	283	AGGGTGAATAATGTACATCCAGGTGAGCGCTGTGTGACTACCTGACCTTCTGCTGCA	342
Qy	41	GluValIlySglGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	343	GAGGTGAAGGAGCAGATTTCAGAGACAGTCCGCCACTCCGGGAACATGACAGGCAATTGAA	402
Qy	61	LeuLeuLeuSerThrLeuGluIlySglYValTrpHisLeuGlyTrpThrArgGluPheVal	80
Db	403	CTGTGCTGAGCACCTTGAGAGAAAGGAGTCTGGCACCTTGTTGACTCGGGAATTCGTG	462
Qy	81	GluAlaLeuArgArgThrGlySerProLeuAlaIlaArgTyrMetAsnProGluLeuThr	100
Db	463	GAGGCCCTCCGAGAACCGGCAGCCCTCTGGCCGCCCTACATGAACCTGAGCTCACG	522
Qy	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	523	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCAATGCTGGAACCTC	582
Qy	121	LeuGlnProThrLeuValAspIlySleuLeuValArgAspValIleuAspIlySlySmetGlu	140
Db	583	CTTCAGCCCACTCTGTGTGACAAAGCTTCTAGTTAGAGACGCTCTTGATTAAGTCATGAG	642
Qy	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	643	GAGGAATGTTGACAAATTGAAGAAGACAAACCGGATTGCTGTCGAGAAAACAATGGAAT	702
Qy	161	GluSerGlyValArgGluLeuLeuIlySArgIleValGlnIlySgluAsnTrpPheSerAla	180
Db	703	GAATCAGGTGTAAAGAGCTACTAAAAAGGATTGTGCAGAAAAGAAAAGTCTCTCTGCA	762
Qy	181	PheLeuAsnValIleuArgGlnThrGlyAsnAsnGluLeuValGlnIlyLeuThrGlySer	200
Db	763	TTTTGAAATGTTCTTCGTCAACAAGAAACAATGAATCTGTCCAAGAGTTAAACAGCTCT	822
Qy	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	823	GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTATCACAAGTTGATGCTCTCAAGTG	882
Qy	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIlySgluValTrpGlyMet	240
Db	883	GAAAGCAACTTCTTTCAACACACAGTTCAAGCCAAATCTGAGAAAGAGGTCTGGGGCATG	942
Qy	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValIleSerGluSerAspThr	260
Db	943	GAGATAACTCATCAGAAATCATCTTTTGACAGATTCTTGTAGTTTCAGAAATCAGACACA	1002
Qy	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
Db	1003	AGTTGGCAGAGAGAGTGTCAAGCTGCTTAGTGAAGTCTTGACATTAACAGCAACATG	1062

QY	281	GLYSerAspSerGlyThrMetGlySerAspSerAspGluAsnValAlaAlaArgAla	300
Db	1063	GGCAGTGAATTCAGGCACCACTGGGAAGTAGATTCAAGTAGAAGAATGTGGCAGCAAGACA	1122
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
Db	1123	TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAAATGGAAAGTTGCCACGACCTTG	1182
QY	321	GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
Db	1183	GAAAGGAAGATATCATCTCTGCCCTCCCTACAGGGAGTGGAAAAAACAGAGTGGCTGTT	1242
QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle	360
Db	1243	TACATTGCCAAGGATCACTTAGACAAGAAGAAAAAGCATCTGAGCCTGAAAAAGTTATA	1302
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380
Db	1303	GTTCTTGTCATAGAAGTACTGCTAGTTGAACAGCTCTTCGCAAGAGTCCAACCAATTT	1362
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1363	TTGAAGAATGTTATCGTTATTGATTAACTGGTGATACCCCACTGAATAATATCATTT	1422
QY	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1423	CCAGAAAGTTGTCAGTCTCTGTGATATTATTATCAGTACAGCTCAAACTCTGAAACTCC	1482
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1483	CTCTTAAACTTGAAAATGGAGAAGATGCTGTGTTCAATTGTCAGACTTTCCCTCATTT	1542
QY	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1543	ATCATGTGATGAATGTCATCAACCAAGAAGACAGTATATAACATCATGAGGCAT	1602
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
Db	1603	TATTTGATGCAGAACTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCATGATTTCC	1662
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500
Db	1663	CTTCCTCAGATACTGGAGTAAACAGCTTCACCTGTGTGGAGGGCCACGAAGCAACC	1722
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1723	AAAGCTGAAGACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACT	1782
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	1783	GTTTAAAGAAAACTTGATCAACTGAAAAAACCAATACAGAGCCATGCCAAGAGTTGCC	1842
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
Db	1843	ATTGCAGATGCAAACGAGAGATCCATTTAAAGAGAACTTTAGAAATPATGACAAGG	1902
QY	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	1903	ATTCAAACTTATTGTCAAATGAGTCCCAATGTCAGATTTTGGAACTCAACCCATGACA	1962
QY	581	TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1963	TGGGCCATTCAAATGGAATAAAAGCTGCAAAAAGAAAGAAATCGCAAGAAGCTGTTGT	2022
QY	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	2023	GCAGAACATTTGAGGAAGTACAATGAGGCCCTACAATAATTATGACAACAATTCGAATG	2082
QY	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal	640
Db	2083	GATCGTATACTCATCTTGAACTTTCTATATATGAAGAGAAAGATTAAGAAGTTTGACGTC	2142

Qy	641	IleGluAspSerAspSerArgLysGlyIleAspAspGluTyrCysAspGlyAspGluAspGlu	660
Dd	2143	ATAGAAGATGATGATGATGAGGGTGATGATGATGATTGTGATGATGATGAAGATGAG	2202
Qy	661	AspAspLeuLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Dd	2203	GATGATTTAAAGAACTTTGAACTGGATGATAACAGATAGATTTCCTGACTTATATT	2262
Qy	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Dd	2263	TTTGAAAACAATTAATGTTGAAAAAGCTGGCTGAAAAACCAGAATATGAAAAAG	2322
Qy	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGlnSerAlaArg	720
Dd	2323	CTGACCAATTAAGAATAACCATATGAGCAATATACTAGACTGAGGAATCAGCACGA	2382
Qy	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu	740
Dd	2383	GGAATATCTTTACAAAAACACGACAGAGTGCATATCGCTTCCAGTGGAATTACTGAA	2442
Qy	741	AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer	760
Dd	2443	AATGAAAAATTGTCTGAAGTAGAGTCAAAGCCCACTCTGATTGGAGCTGGACACAGC	2502
Qy	761	SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr	780
Dd	2503	AGTGAGTTCAAACCATGACACAGAAATGAACAAAAAAGATCATTAATAATTTCGCACT	2562
Qy	781	GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu	800
Dd	2563	GGAAAAATAATCTGCTTATCGCTACCACAGTGGCAGAAGAAGCTGTGATATTAAAGAA	2622
Qy	801	CysAsnIleValIleArgTyrGlyLeuValThraSngluIleAlaMetValGlnAlaArg	820
Dd	2623	TGTAACTATTGTTATCCGTTATGTTCTCTCTACCAATGAATAGCCATGTGTCAGGCCCGT	2682
Qy	821	GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal	840
Dd	2683	GGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTGTCTCACAGTGGTTCAAGAGTT	2742
Qy	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860
Dd	2743	ATCGAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGATATAAGCTATACTTGT	2802
Qy	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer	880
Dd	2803	GTTCAAAATATGAAACCAAGAGATATGCTCATTAAGATTTTGAATTACAGATGCCAAAGT	2862
Qy	881	IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro	900
Dd	2863	ATAATGAAAAAGAAATGAAAAACCAAGAGAATATTGCCAAGCATTACAAGAATTAACCA	2922
Qy	901	SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle	920
Dd	2923	TCACTAATACTTCTTTGCAAAAACTGCAGCTGTCTAGCCGTCTTGGGGAAGATATC	2982
Qy	921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle	940
Dd	2983	CATGTAATTGAGAAAAATGCATCACGTCATATATGACCCAGAAATTCAGGAACCTTACATT	3042
Qy	941	ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle	960
Dd	3043	GTAAGAGAAAAACAACCACTGCAAAAAGAGTGCCGACTATCAATATAATGTGAAATC	3102
Qy	961	IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro	980
Dd	3103	ATCTGCAATATGTGGCCAGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGGCT	3162
Qy	981	CysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnTyr	1000
Dd	3163	TGTCCTCAAAATAAGCAATTTTGTAGTGTTTTCAAAAAATTAATTCAACAAGAAACAATAC	3222
Qy	1001	LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu	1020

|||||  
Db 3223 AAAAAGTGGGTAGAAATTACCTATCAACATTTCCCAATCTTGACTATTCAAGATGCTGTTA 3282  
Qy 1021 PheSerAspGluAsp 1025  
|||||  
Db 3283 TTTAGTGATGAGGAT 3297  
  
RESULT 7  
CS042815  
LOCUS CS042815 3434 bp DNA linear PAT 22-MAR-2005  
DEFINITION Sequence 3369 from Patent WO2005019258.  
ACCESSION CS042815  
VERSION CS042815.1 GI:61849927  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
  
REFERENCE  
1  
AUTHORS Abbas, A., Clark, H., Ouyang, W., Williams, P.M., Wood, W.I. and Wu, T.D.  
TITLE Compositions and methods for the treatment of immune related  
diseases  
JOURNAL Patent: WO 2005019258-A 3369 03-MAR-2005;  
Genentech, Inc. (US)  
FEATURES  
Source Location/Qualifiers  
1..3434  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
  
ORIGIN  
  
Alignment Scores:  
Pred. No.: 3.14e-312 Length: 3434  
Score: 5307.00 Matches: 1024  
Percent Similarity: 100.0% Conservative: 1  
Best Local Similarity: 99.9% Mismatches: 0  
Query Match: 99.9% Indels: 0  
Gaps: 0  
DB: 6  
  
US-09-515-363C-2 (1-1025) x CS042815 (1-3434)  
  
Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
Db 223 ATGTGCAATGGGTATTCCACAGACGAAGATTCCGCTATCTCATCTCGTGTTCAGGGCC 282  
Qy 21 ArgValIysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
Db 283 AGGGTGAAGAAATGATACATCCAGGTGGAAGCCTGTGCTGACTACCTGACCTTTCGCTGCA 342  
Qy 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 343 GAGGTGAAGAGCAGATTCCAGAGCAGAGTCCGCACTCCGGGAACATGCAGGCAATTGAA 402  
Qy 61 LeuLeuLeuSerThrLeuGluIysGlyValTyrPheIleGlyTyrThrArgGluPheVal 80  
Db 403 CTGCTGCTGAGCACCTTGAGAGGAGGTCTGGCACCTTGCTTGGACTCGGGAATTCTGTG 462  
Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
Db 463 GAGGCCCTCCGAGAACCGGCACCTCTGCGCCGCTTAATGAAACCTGAGCTCAGC 522  
Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 523 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 582  
Qy 121 LeuGlnProThrLeuValAspIysLeuLeuValArgAspValLeuAspIysCysMetGlu 140  
Db 583 CTTCAAGCCCACTCTGTGAGACACTTCTAGTAGACGTCTTGATTAAGTGCAAGAG 642  
Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
Db 643 GAGGAAGTGTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAACAATGGAAT 702

Qy 161 GluSerGlyValArgGluLeuLeuIysArgIleValGlnIysGluAsnTrpPheSerAla 180  
Db 703 GAATCAGGTGTAAAGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAAGTGTCTCTGCA 762  
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 763 TTTCTGAATGTTCTTCGTCAAAACAGAAAACATGAACCTTGTCCAAGAGTTAAACAGGCTCT 822  
Qy 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 823 GATTGCTCAGAAAGCATGCAGAGATTGAGATTATCACAGTTGATGGTCTCAAGTG 882  
Qy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIysGluValTrpGlyMet 240  
Db 883 GAAGAGCAACTTCTTCAACACACAGTTCCAGCCAAATCTGGAGAGAGGCTCTGGGGCATG 942  
Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
Db 943 GAGAAATACTCATCAGAAATCATCTTTGAGATTCTTCTGTAGTTTCAGAAATCAGACACA 1002  
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 1003 AGTTTGGCAGAAAGAGTGTACGCTGCTTAGATGAAGAACTCTTGACATTAACAGCAACATG 1062  
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
Db 1063 GGCAGTGATTCAGGCACCACTGGGAAGTGATTCAGATGAAGAAATGTGGCAGCAAGACA 1122  
Qy 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
Db 1123 TCCCCGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCCAAGCCTTG 1182  
Qy 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340  
Db 1183 GAAGGGAAGAATATCATCATCTGCTCCCTACAGGAGTGGAAAAACCAAGTGGCTGTT 1242  
Qy 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360  
Db 1243 TACATTGCCAAGATCATTAGACAAGAAAAGCATCTGAGCCTGGAAAAAGTTATA 1302  
Qy 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380  
Db 1303 GTTCTTGTCATAAAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAACCATTT 1362  
Qy 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400  
Db 1363 TTGAAGAAATGTATCGTGTATTGATTAGTGTGATACCCAACTGAAATATCATATT 1422  
Qy 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
Db 1423 CCAGAAAGTTGTCAAGTCTGTGATATTATTATCATGATACAGCTCAAAATCTTGAAAACTCC 1482  
Qy 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
Db 1483 CTCTTAAACTTGMAAATGAGAAAGATGCTGTGTTCAATTGTCAAGACTTTTCCCTCATTT 1542  
Qy 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460  
Db 1543 ATCATTGATGAATGTATCATCACCAACAAAGAACAGTGTATAATACATCATGAGGCAT 1602  
Qy 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysProValIlePro 480  
Db 1603 TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAACAACCAAGTGAATCCC 1662  
Qy 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAla 500  
Db 1663 CTTCTCAGATACTGGAGCTTAACAGCTTCAACCTGTGTTGGAGGGGCCCAAGCAAGCC 1722  
Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
Db 1723 AAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATGATTTACTATTAAAACT 1782

QY	521	VallYsGIuAsnLeuAspGIuLeuLysAsnGIuNIeGIuGIuProCysLysLysPheAla	540
Db	1783	GTTAAAGAAAACCTTGATCAACTGAAAAAACCAATAACAGAGCCATGCAGAAAGTTTGCC	1842
QY	541	IlEAlAspAlaThrArgGIuAspProPheLysGIuLysLeuGIuNIeMetThrArg	560
Db	1843	ATTGCAGATGCAACCCAGAGAAGATCCATTAAAGAGAACTTCTAGAAATAATGACAAG	1902
QY	561	IlEgInThrTyrCysGIuMetSerPrometSerAspPheGIuThrGIuProTyrGIuGIu	580
Db	1903	ATTCAACTTATTTGTCAATAGAGTCCAAATGTCAAGATTTTGGAACCTCAACCTATGACAA	1962
QY	581	TRPAlaIlEgImetCylLysLysAlaAlaLysLysGIuAsnArgLysGIuArgValCys	600
Db	1963	TGGGCATTCAATGGAAAAAAAGCTGCAAAAGAGGAATCGCAAGAACGTGTTGT	2022
QY	601	AlaGIuHIslEuarGlysTyrAsnGIuAlaLeuGIuNIeAsnAspThrIlEargMetIlE	620
Db	2023	GCAGACATTTGAGGAGTACATGAGGCCCTCAAAATTAATGACACAAATTGCAATGATA	2082
QY	621	AspAlaTyrThrHisLeuGIuThrPheTyrAsnGIuLysAspLysLysPheAlaVal	640
Db	2083	GATGCGTATCTCATCTTGAAACTTTCTATATGAGAGAAAGATTAAGAGTTTGCACTC	2142
QY	641	IlEgIuAspAspSerAspGIuGIuGIuYAspAspGIuTyrCysAspGIuAspGIu	660
Db	2143	ATGAGAGATGATAGTGAAGGGGTGATGATGATGATGATGATGATGATGATGATGAG	2202
QY	661	AspAspLeuLysLysProLeuLysLeuAspGIuThrAspArgPheLeuMetThrLeuPhe	680
Db	2203	GATGATTTAAAGAAACCTTGAAACTGATGAAACAGATAGATTTCTCATGACTTATTT	2262
QY	681	PheGIuAsnAsnLysMetLeuLysArgLeuAlaGIuAsnProGIuTyrGIuAsnGIuLys	700
Db	2263	TTTGAACAACATATAATGTTGAAGAGGCTGGCTGAAAAACCAAGATATGAAAAATGAAAAAG	2322
QY	701	LeuThrLysLeuArgAsnThrIlEmetGIuGIuTyrThrArgThrGIuGIuSerAlaArg	720
Db	2323	CTGACCAATTAAGAAATACCATATGAGCAATATTAAGACTGAGGAATGACACGA	2382
QY	721	GIuTIlEIlEpherThrLysThrArgGIuSerAlaTyrAlaLeuSerGIuTrpIlEThrGIu	740
Db	2383	GGAATAATCTTTACAAAAACAGACAGAGTGCATGTGCGCTTCCAGTGGATTACTGAA	2442
QY	741	AsnGIuLysPheAlaGIuValGIuValLysAlaHisLysLeuIlEgIuAlaGIuHisSer	760
Db	2443	AATGAAAAATTTGCTGAAGTAGAGTCAAGCCCACTCTGATTGGAGCTGGACACAGC	2502
QY	761	SerGIuPheLysProMetThrGIuAsnGIuGIuLysGIuValIlEserLysPheArgThr	780
Db	2503	AGTGAGTTCMAACCCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAATTTCCACT	2562
QY	781	GIuLysIlEAsnLeuLeuIlEAlaThrThrValaGIuGIuLysAspIlELeysGIu	800
Db	2563	GGAAAAATAATCTGCTTATCGCTACACACAGTGGCAGAGAAGGTCTGATATTAAAGAA	2622
QY	801	CysAsnIlEValIlEArgTyrGIuLeuValThrAsnGIuIlEAlaMetValGIuAlaArg	820
Db	2623	TGTAACATTGTTATCCGTTATGCTCTGTCACCAATGAATAAGCCATGGTCCAGGCCCGT	2682
QY	821	GIuArgAlaArgAlaAspGIuSerThrTyrValLeuValAlaHisSerGIuSerGIuVal	840
Db	2683	GGTCGAGCCAGAGCTGATGAGACACCTACGTCCTGTGCTCACAGTGGTTCAGAGTT	2742
QY	841	IlEgIuHIslGIuThrValaAsnAspPheArgGIuLysMetMetTyrLysAlaIlEHisCys	860
Db	2743	ATCGAACATGAGACAGTTAATGATTTCCGAGAGAGATGATGATAAAGCTATACATTGT	2802
QY	861	ValGIuAsnMetLysProGIuGIuTyrAlaHisLysIlELeuGIuLeuGIuMetGIuSer	880
Db	2803	GTTCAAAATATGAAACCCAGAGAGTATGCTCATTAAGATTTTGGAAATTACAGATGCAAGT	2862
QY	881	IlEmetGIuLysLysMetLysThrLysArgAsnIlEAlaLysHisTyrLysAsnAsnPro	900

Db	2863	ATAATGAAAAGAAAATGAAAAACCAAGAGAAATATTGCCAAGCATTTACAAGATAACCCA	2922
Qy	901	SerLeuIleThrPheLeuCysIlyAsnCyseSerValLeuAlaCysSerGlyIleAspIle	920
Db	2923	TCACATAATACTTCTCTTGCAAAAACTGCAGCTGTGCTAGCCGTGTTCTGGGGAGATATC	2982
Qy	921	HisValIleGlyIlyMetHisHisValAsnMetThrProGluPheIlyGluLeuTyrIle	940
Db	2983	CATGTAATTGAGAAATGCATCAGCTCAATATGACCCAGAAATTCAGGAACTTTACATT	3042
Qy	941	ValArgGluAsnIlySAlaLeuGlnIlySlyCysAlaAspTyrGlnIleAsnGlyIle	960
Db	3043	GTAAGAGAAAACAAGCACTGCAAAAGAGTGTGCCGACTATCAATATAATGGTGAATA	3102
Qy	961	IleCysIlyCysGlyGlnAlaTyrGlyThrMetMetValHisIlySglyLeuAspLeuPro	980
Db	3103	ATCTGCAAAATGTGCCAGGCTTGCGGAGACAATGATGCTGCACAAAGGCTTAGATTGCC	3162
Qy	981	CysLeuIlySAlaArgAsnPheValValAlaPheIlySAsnAsnSerThrIlySglyTyr	1000
Db	3163	TGTCCTCAAAATAGGAATTTTGTAGTGGTTTCAAAAATTAATTCACAAAGAAACAATAC	3222
Qy	1001	LysIlyStryValGluLeuProIleThrPheProAsnLeuAspTyrSerGlyCysLeu	1020
Db	3223	AAAAAGTGGGTAGAAATTAACCTATCACATTTCCCAATCTTGACTATTCAAGATGCTGTTTA	3282
Qy	1021	PheSerAspGluAsp	1025
Db	3283	TTTAGTGATGAGGAT	3297
RESULT	8		
LOCUS	CS045628	3434 bp	DNA linear PAT 22-MAR-2005
DEFINITION	Sequence 6182 from Patent WO2005019258.		
ACCESSION	CS045628		
VERSION	CS045628.1	GI:61851939	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Abbas, A., Clark, H., Ouyang, W., Williams, P. M., Wood, W. I. and Wu, T. D.		
TITLE	Compositions and methods for the treatment of immune related diseases		
JOURNAL	Patent: WO 2005019258-A 6182 03-MAR-2005;		
Genentech, Inc. (US)			
FEATURES	Location/Qualifiers		
source	1..3434		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:			
Pred. No.:	3.14e-312	Length:	3434
Score:	5307.00	Matches:	1024
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	99.9%	Mismatches:	0
Query Match:	99.9%	Indels:	0
DB:	6	Gaps:	0
US-09-515-363C-2 (1-1025) x CS045628 (1-3434)			
Qy	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	223	ATGTGCAATGGGTATTCACACAGACGAGAAATTCGCTATCTCATCTCGTCTTCAGGGCC	282
Qy	21	ArgValIysMetTyrIleGlnValGluProValLeuAspTyrIleuThrPheLeuProAla	40
Db	283	AGGCTGAAAATGTACATCCAGGTGAGCCTGTGCTGAGACTACCTGACCTTCTGCCTGCA	342

QY	41	GIUVallySGlUGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	343	GAGGTGAAGAGACAGATTCAAGAGACAGTCGCCACTCCGGAAACATGACAGCAGTTGAA	402
QY	61	LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal	80
Db	403	CTGCTGCTGAGCACCTTGAGAGAGGAGTCTGGCACCTTGTTGACTCGGAAATTCCGTG	462
QY	81	GIUAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
Db	463	GAGGCCCTCCGAGAACCCGACCCCTCTGGCCGCCCTTACATGAACCTGAGCTCACG	522
QY	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	523	GACTTGCCCTCTCCATCGTTTGAAACGCTCATGATGAATATCTCCAACGTGCTGAACCTC	582
QY	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
Db	583	CTTCAGCCCACTCTGGTGACAGCTTCTAGTTAGACGCTCTTGATTAAGTGCAATGAG	642
QY	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	643	GAGGAACGTGTGCAATTGAAGACAGAAACCCGATTGCTGCTGCAGAAACCAATGGAAT	702
QY	161	GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla	180
Db	703	GAAATCAGGTGTAAAGAGACTAATAAAAGATTGTTCAGAAAGAAAACCTGTTCTTCGA	762
QY	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	763	TTTCTGAATGTTCTTCGTCAACAGGAAACAATGAACCTTGTCCAAAGATTAAACAGCTCT	822
QY	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	823	GATTGCTCAGAAAGCAATGCAGAGATTGAGAATTATCACAAGTTGATGTCCTCAAGTG	882
QY	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240
Db	883	GAAAGCAACTCTTTCAACCAACAGTTCAGCCAAATCTGGAGAAAGAGTCTGGGGCATG	942
QY	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260
Db	943	GAGAATACTCATCAGAAATCATCTTTGCAGATTCTTCTAGATTTCAGAAATCAGACACA	1002
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
Db	1003	AGTTTGGCAGAAAGAGTGTCAAGCTGCTTAGATGAAGTCTTGACATTAACAGCAACATG	1062
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	1063	GGCAGTGATTCAGGCACCATGGGAAGTGATTCAGATGAAGAGAATGTGGCAGCAAGACA	1122
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
Db	1123	TCCCCGGAGCCAGAACTCCAGCTCAAGCCTTACCAAATGGAAAGTTGCCCAAGCCAGCTTG	1182
QY	321	GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
Db	1183	GAAAGGAAGAATATCATCATCTGCTCCCTACAGGAGAGTGAAAAAACAGAGTGGCTGTT	1242
QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysLysLysAlaSerGluProGlyLysValIle	360
Db	1243	TACATTGCCAAGGATCACTTGACAAAGAAAAAAGCACTGAGCCTGAAAAAGTTATA	1302
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380
Db	1303	GTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGATTCCCAACCATTT	1362
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1363	TTGAAGAAATGGTATCGTGTATTGGAATTAAAGTGTGATACCCCAACTGAAAAATATCATTT	1422

QY	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1423	CCAGAA GTTGTCAAGTCTCTGTGATATTATTATCAGTACAGCTCAATCTTGAAAACTCC	1482
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1483	CTCTTAACTTGAAAAATGGAGAAGATGCTGGTGTCAATTGTCAAGCTTTCCCTCATT	1542
QY	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1543	ATCATGTGATGATGTCTATCATCACCAACAAGAGCAGTGTATATATACATCATGAGGCAT	1602
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysProValIlePro	480
Db	1603	TATTGTGATGCAGAGTTGAAAAACAATAGACTCAAGAAAAACAACCAAGTATTCCC	1662
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyValAlaThrLysGlnAla	500
Db	1663	CTTCCTCAGATACTGGGACTTAACAGCTTCACTGGTGTGGAGGGCCAGCAAGCACC	1722
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1723	AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTTAAACT	1782
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysPheAla	540
Db	1783	GTTAAAGAAAACTTGATCAACTGAAAAACAATAACAGAGCCATGCAAGATTGCC	1842
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGlnIleMetThrArg	560
Db	1843	ATTGCAAGTGCACACGAGAGATCCATTTAAAGAGAACTTCTAGAAATAATGACAAGG	1902
QY	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	1903	ATTCAAACTTATTGTCAAAATGAGTCCAAATGTGCAGATTTTGGAACTCAACCTATGACAA	1962
QY	581	TyrAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1963	TGGGCATTTCAATGAAAAAAAAGCTGCAAAAGAGAAATCGCAAGAACGTGTTGT	2022
QY	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	2023	GCAGAACATTTGAGAGATGACATGAGGCCCTACAATAATGACACAAATTCGAATGATA	2082
QY	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAlaVal	640
Db	2083	GATGCCGATATCTCATCTTGAACCTTCTATATATGAGAAAGATTAAGAACTTGCAGTC	2142
QY	641	IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	2143	ATAGAAATGATAGTAGAGGCTGTGATGATGAGTATTGTGATGGTGAAGAAGATGAG	2202
QY	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	2203	GATGATTTAAAGAAACCTTTGAAACTGGATGAACAGATAGATTTCTCATGACTTTATTT	2262
QY	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db	2263	TTTGAACAACAATAAATGTTGAAAAAGCGCTGCTGAAAAACCCAGATATGAAAAAG	2322
QY	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg	720
Db	2323	CTGACCAAAATTAAGAAATACCAATATGAGCAATATACTAGGACTGAGGAATCAACACGA	2382
QY	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGlu	740
Db	2383	GGAATAATCTTACAAAAACAGACAGAGTGCAATATGCGCTTCCCACTGCAATTACTGAA	2442
QY	741	AsnGluLysPheAlaGluValGlyValLysAlaHisHisIleuIleGlyValaGlyHisSer	760
Db	2443	AATGAAAAATTTGCTGAAGTAGAGATCAAAAGCCCACTCTGATTGGAGCTGCACACAGC	2502
QY	761	SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr	780

Db	2503	AGTGAGTTCAAACCCATGACACAGATGACAAAGAGTCACTAGTAATTTCCGACT	2562
Qy	781	GlyLysIleAsnLeuLeuIleAlaThrValAlaGluGluGlyLeuAspIleLysGlu	800
Db	2563	GGAAAAATAATCTGCTTATCGCTACACAGTGGCAGAGAAGGCTCGATATTAAAGAA	2622
Qy	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	820
Db	2623	TGTAACTATTGTTATCCGTTATGGTCTCGTCAACCAATGAATAGCCATGGTCCAGGCCGT	2682
Qy	821	GlyArgAlaArgAlaAspGluSerThrTyrValIleValAlaHisSerGlySerGlyVal	840
Db	2683	GGTCGAGCCAGAGCTGATGAGACACTACGTCCTGGTGTCCACAGTGGTTCAGAGATT	2742
Qy	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860
Db	2743	ATCGAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATATAAGCTATACATTGT	2802
Qy	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer	880
Db	2803	GTTCAAAATATGAAACCAAGAGAGTATGCTCATAGATTTTGAATTACAGATGCAAAAGT	2862
Qy	881	IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro	900
Db	2863	ATAATGAAAAAGAAATGAAACCAAGAAATATGGCAAGCATTACAAAGATAAACCA	2922
Qy	901	SerLeuIleThrPheLeuCysLysAsnCysSerValIleuAlaCysSerGlyGluAspIle	920
Db	2923	TCACTAATAACTTTCCTTGGCAAAAACTGCAGTGTGTAGCTGTCTGGGGAAGATATC	2982
Qy	921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle	940
Db	2983	CATGTAATTGAGAAATATGATCATCGTCAATATGACCCAGAATTCAAGGAATTACATT	3042
Qy	941	ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGluIle	960
Db	3043	GTAAGAGAAAAACAAGCAGCTGCAAAAAGAGTGTGCCACTATCAATAATGTGAAATC	3102
Qy	961	IleCysLysCysGlyGlnAlaIleTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro	980
Db	3103	ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTTGCCT	3162
Qy	981	CysLeuLysIleArgAsnPheValValAlaPheLysAsnAsnSerThrLysLysGlnTyr	1000
Db	3163	TGTCTCAAAATAAGGAATTTTGTAGTGTTCCTTCAAAATAAATTCAACAAGAAACAATAC	3222
Qy	1001	LysLeuTyrPheValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLeu	1020
Db	3223	AAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTGAGAATGCTGTTTA	3282
Qy	1021	PheSerAspGluAsp 1025	
Db	3283	TTTAGTGATGAGGAT 3297	

RESULT 9  
AX300832 3372 bp DNA linear PAT 16-APR-2005  
LOCUS  
DEFINITION Sequence 1 from Patent WO0185955.  
ACCESSION AX300832  
VERSION AX300832.1 GI:17382110

KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.

REFERENCE  
AUTHORS Bahr, G., Cocude, C. and Capron, A.  
TITLE Rh16 polypeptides and its fragments and polynucleotides encoding  
said polypeptides and therapeutic uses  
JOURNAL Patent: WO 0185955-A 1 15-NOV-2001;  
Istac (FR) ; INSTITUT PASTEUR DE LILLE (FR)

FEATURES		Location/Qualifiers
source		1..3372 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 155..3332 /note="unnamed protein product"
CDS		/codon_start=1 /protein_id="CAD13049.1" /db_xref="GI:17382111" /translation="MSGYSTDENFRYLISCFRAVKMYIQVEPVLDTLTFLLPAEVEKE QIQRTVATSGNMQAVELLSTLEKGVWHLGWTREVEALRRTGSPLAARYMNPETDL PSPFENAHDEVLQLLNLQPTLVKLLVRDVLDKMEEBELTIEDRNRIAAENNGN ESGVRELKRIYQKENWFSALFNVLVQTNNELVQELTGSSESNAEINLSQVDP QVEBQLSTVQPNLEKEVGMENNSSESSPADSSVSESDTSLAEGSVCLDESIGH NSNMGSDSGTMSDSDSEENVAAARASPEPELQRPYOMEVAQPALEGKNIICLPYSG KTRVAVYIAKDHLDKKKAASEPGKVIYLVNKLVLVEQLFRKFPQPLKMYRVIGLSG DTQLKISFEVYVSCDIIISTAQILNSLNLNENGEDAGVQSDSFIIDECHTNK EAVYNNIMRHYLMQKLKNRLLKKNRPVILPQILGLTASPGVGATKQAKAEHILK LCANLDAFTITVKENLDQKNOIQEPCCKFAIADATREDPFREKLEIMTRIQTVCQ MSPMSDFGTOPYEQWAIQMEKKAAGNRKESVCAEHLRKYNKALQINDTIRMIDAYT HLETFYNEEKDKKFAVIDEDSDSGDEYCDGDEDEDDLKPKLDETDRFLMTLFE NNKMLKRLAENDEYENEKLTJLKRNTIMEQYTRTEESARGIIFPKTRQSAVALSOWITE NEKFAEVGVAAHLIGAGHSSEFKPTQNEQKEVISKFRGKINLLIATVAEEGLDI KECNIVIRGLVTNEIAMVQARGRADESTYVLVAHSGSGVIRETVNDFREKMMYK AHCVNQMKPEBYAHKILELOMOSINEKMKTKRNIAKHYKNPSLITFLCKNCSVLA CSGEDIHVIEKMHVNMTPBEFKELYVRENKALQKKCADYQNGEIIICKGOAWGTMM VHKGLDLPCLKIRNFVVFKNNSTKYKKWVELPITFPNLDYSECCLPFSD"
ORIGIN		
Alignment Scores:		
Pred. No.:		6.64e-311 3372
Score:		5285.00 Matches: 1020
Percent Similarity:		99.7% Conservative: 2
Best Local Similarity:		99.5% Mismatches: 3
Query Match:		99.5% Indels: 0
DB:		6 Gaps: 0
US-09-515-363C-2 (1-1025) x AX300832 (1-3372)		
Qy	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
Db	155	ATGTGGAATGGGTATTCACAGACGAGAATTTCCGCTATCTCATCTCGTCTTCAGGGCC 214
Qy	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
Db	215	AGGGTGAATAATGTACATCCAGGTGAGCCTGTGCTGGACTACCTGACCTTTCTGCCTGCA 274
Qy	41	GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
Db	275	GAGGTGAAGGAGCAGATTGAGAGCAGTGGCCACTTCGGGAACATGACGAGGAGTTGAA 334
Qy	61	LeuLeuLeuSerThrLeuGluLysGlyValIleTyrHisLeuGlyTyrThrArgGluPheVal 80
Db	335	CTGCTGCTGAGCACCTTGAGAAAGGAGTCTGGCACTTGTTGACTCGGAAATTGCTG 394
Qy	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
Db	395	GAGGCCCTCCGAGAACGGCAGCCCTCTGGCGCCGCTACATGAACCTGAGACTCAGC 454
Qy	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
Db	455	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 514
Qy	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
Db	515	CTTGAGCCCACTCTGTGTGACAAAGCTTCTAGTTAGAGACGCTCTTGATTAAGTCATGGAG 574
Qy	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
Db	575	GAGGAAGTGTGACAAATGAAGACAGAAACCGGATTGCTGCTGCAGAAAACAATGGAAT 634
Qy	161	GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180

Db 635 GAATCAGGTGTAGAGAGCTAAATAAGGATGTGTGAGAAAAGAAACGTGTTCTCTGCA 694  
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 695 TTTCTGAATGTCTCTCGTCAACAGGAAACAATGAATCTGTCCAAGAGTTAACAGGCTCT 754  
Qy 201 AspCysSerGluSerAspAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 755 GATTGCTCAGAAAGCAATGCAGAGATTGAGATTTCACAAGTTGATGGTCTCAAGTG 814  
Qy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrrGlyMet 240  
Db 815 GAAGAGCAACTCTTTCACACCACAGTTCAGCCAAATCTGGAGAAAGAGGTCTGGGGCATG 874  
Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
Db 875 GAGATAAATCATCAGAAATCACTTTTGACAGATCTTCTGTAGTTTCAGAAATCAGACACA 934  
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 935 AGTTGGCAGAGAGAGAGTGTCACTGCTTAGATGAAGTCTTGACATAACAGCAACATG 994  
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
Db 995 GGCAGTGATTCAGGCCACCATGGGAAGTGATTCAAGATGAAGAAATGTGGCCAGACAAGCA 1054  
Qy 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
Db 1055 TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAAATGGAAAGTTGCCCGACGCTTG 1114  
Qy 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340  
Db 1115 GAAGGAAAGAATATCATCATCTGCCTCCCTACAGGGAGTGAAGAAAAACAGAGTGGCTGTT 1174  
Qy 341 TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle 360  
Db 1175 TACATTGCCAAGGATCATTAGACAAGAAAAAGCATCTGAGCCTGGAAGATTATA 1234  
Qy 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380  
Db 1235 GTTCTTGTCATAAGTACTGCTAGTGAACAGCTCTCCGCAAGAGTTCCAAACCATTT 1294  
Qy 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400  
Db 1295 TTGAAGAAATGCTATCGTGTATTGGATTAAAGTGTGATACCCCACTGAAATATCATTT 1354  
Qy 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
Db 1355 CCAGAAATTGTCAAGTCCGTGTATTATTATCAGTACAGCTCAAAATCTTGAAAACTCC 1414  
Qy 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
Db 1415 CTCTTAAACTTGAAAAATGAGAGAGATGCTGTGTTCAATTGTCAGACTTTCTTCATT 1474  
Qy 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460  
Db 1475 ATCATTTGATGATGTCTATCACACCAACAAGAGCAGTGTATAATAACATCATGAGGCAT 1534  
Qy 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480  
Db 1535 TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATCCC 1594  
Qy 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500  
Db 1595 CTTCTTCAGATTACTGGGACTTAACAGCTTCACTGGTGTGGAGGGGCCACGAGCAAGCC 1654  
Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
Db 1655 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCAATTTACTATTAAACT 1714  
Qy 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540

Db 1715 GTTAAAGAAAACCTTGATCAACTGAAAAACCAATATACAGAGCCATGCAGAAAGTTGCC 1774  
Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560  
Db 1775 ATTGCAGATGCACACCAGAGAAAGATCCATTTTAAAGAAAACTTCTAGAAATATGACAAG 1834  
Qy 561 IleGlnThrTyrCysGlnMetSerPrometSerAspPheGlyThrGlnProTyrGluGln 580  
Db 1835 ATTCAACTTATTTGTCAAAATGAGTCCAAATGTACAGATTTTGAACCTCAACCTTATGAACA 1894  
Qy 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
Db 1895 TGGGCCATTCAATGGAAGAAAAAGCTGCAAAAGAAAGAAATCGCAAGAAAGTGTGT 1954  
Qy 601 AlaGlnHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
Db 1955 GCAGAACATTTGAGGAAGTACAATAAGGCCCTACAAATTAATGACACAAATTCGAATGATA 2014  
Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysAspLysLysPheAlaVal 640  
Db 2015 GATGCCATTAATCATCTTGAACCTTCTATATGAAGAAAGATAGAGATTGCAAGTTC 2074  
Qy 641 IleGluAspAspSerAspGluGlyLysAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 2075 ATAGAGATGATAGTAGTAGGGGTGCTGATGATGAGTATTGTGATGGTGTGAAGATGAG 2134  
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2135 GATGATTTAAAGAAACCTTGAACCTGATGAACAGATAGATTCTCATGCTTTATTT 2194  
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2195 TTTGAAACAATAAATGTTGAAAAAGGCTGCTGTAAGAACCAAGATATGAATAAGAAAG 2254  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720  
Db 2255 CTGACCAATTAAGAAATACATAATGAGCAATATACAGACATGAGAAATCAGACAGA 2314  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740  
Db 2315 GGAATATCTTTACAAAAACAGACAGAGTGCATATGCCGTTTCCAGTGGATTACTGAA 2374  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisIleLeuIleGlyAlaGlyHisSer 760  
Db 2375 AATGAAAAATTGCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGAACACAGC 2434  
Qy 761 SerGluPheLysPrometThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2435 AGTGAGTTCAAACCCATGACACAGAATGAACAAAAAGAGTCATTAGTAAATTTCCCACT 2494  
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2495 GGAATAATAATCTGCTTATGCTACACAGTGGCAGAAAGAGTGTGATATTAAAGAA 2554  
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2555 TGTAACTTTGTTATCCGTTATGCTCTGTCACCAATGAATAAGCATGTGCCAGCCGT 2614  
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2615 GGTGAGCCAGAGCTGATGAGACACCTACGTCTGTGTTGCTCACAGTGGTTCAGAGATT 2674  
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860  
Db 2675 ATCGAACGTGAGACAGTTAATGATTTCCGAGAGAAAGATGATATMAAGCTATATCATTTGT 2734  
Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2735 GTTCAAAATATGAACAACAGAGAGTATGCTCATAGATTTTGAATTTACAGATTCADAAGT 2794  
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900  
Db 2795 ATATGGAAGAAAGAAATGAAGAACCAAGAGAAATATGCAAGCATTAACAAGATAACCCA 2854

QY 901 SerleuileThrphelucySlysAsnCysSerValleuAlaCysSerGIyGIuAspIle 920  
| | | | |  
Db 2855 TCACTAATACTTCCITTTGCCAAAACTGCAGTGTGCTAGCCCTGTTCTGGGGAAGATATC 2914  
QY 921 HisValIleGIuLysMetHisHisValAsnMetThrProGIuPhelysGIuLeuTyrIle 940  
| | | | |  
Db 2915 CATGTAATTGAGAAAATGATCATCGTCAATATGACCCAGAAATTCAGGAACCTTTACATT 2974  
QY 941 ValArgGIuAsnLysAlaLeuGIuLysLysCysAlaAspTyrGIuLysAsnGIyGIuIle 960  
| | | | |  
Db 2975 GTAAGAGAAAAACAAGACACTGCAGAAAGAGTGTGCCGACTATCAAAATAATGGTGAATC 3034  
QY 961 IleCysLysCysGIyGIuAlaTrrpGIyThrMetMetValHisLysGIyLeuAspLeuPro 980  
| | | | |  
Db 3035 ATCTGCAATGTGGCCAGCGCTGGGGAACAATGATGTGCACAAAGGCTTAGATTGGCCT 3094  
QY 981 CysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGIuTyr 1000  
| | | | |  
Db 3095 TGTCTCAAAATAAGGAATTTTGTAGTGTTCAAAAATAATTCAACAAGAAACAATAC 3154  
QY 1001 LysLysTrrpValGIuLeuProIleThrPheProAsnLeuAspTyrSerGIuLysCysLys 1020  
| | | | |  
Db 3155 AAAAAGTGGTAGAATTACTATCATATTCCCAATCTTGACTATTGAGAAATGCTGTTA 3214  
QY 1021 PheSerAspGIuAsp 1025  
| | | | |  
Db 3215 TTTAGTGATGAGGAT 3229  
RESULT 10  
AY017378 3373 bp mRNA linear PRI 28-NOV-2003  
LOCUS Homo sapiens RNA helicase-DEAD box protein RH116 mRNA, complete  
DEFINITION cds.  
ACCESSION AY017378 GI:12621065  
VERSION AY017378.1  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 3373)  
AUTHORS Cocude,C., Truong,M.-J., Billaut-Mulot,O., Delsart,V.,  
Darcissac,E., Capron,A., Mouton,Y. and Bahr,G.M.  
TITLE A novel cellular RNA helicase, RH116, differentially regulates cell  
growth, programmed cell death and human immunodeficiency virus type  
1 replication  
J. Gen. Virol. 84 (12), 3215-3225 (2003)  
JOURNAL 2 (bases 1 to 3373)  
REFERENCE Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J.,  
AUTHORS Capron,A. and Bahr,G.M.  
TITLE Identification of a new RNA helicase (RH116) regulated by the  
immunomodulator Murabutide  
Unpublished  
JOURNAL 3 (bases 1 to 3373)  
REFERENCE Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J.,  
AUTHORS Capron,A. and Bahr,G.M.  
TITLE Direct Submission  
JOURNAL Submitted (09-JAN-2001) Laboratoire d'Immunologie Molculaire de  
l'infection et de l'inflammation, Institut Pasteur de Lille, 1 rue  
du Professeur Calmette BP 245, Lille 59019, France  
FEATURES  
source 1..3373  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
155..3232  
/codon\_start=1  
/product="RNA helicase-DEAD box protein RH116"  
/protein\_id="AAG54076.1"  
/db\_xref="GI:12621066"  
/translation="MSGYSTDENFRYLISCFRARVKMYIQVEPVLDTLFLPAEVKE

QIORTVATSGNMQAVELLSTLEKGVHLGWTREFVEALRRTGSPLAARYMPELTDL  
PSPSFENAHDEYLQNLNLQPTLVLDKLVLDVLDKCMSEBELTIEDNRRIAAENNGN  
ESGVELLKRIVOKENWSPAFLNLVLRQTNNELVQELTGSDCSSNAIEMLSQVDGP  
QVEBQLSTYQPNLEKEVWGMENSSSESSPADSSVSESDPSLAEGSVSLDESIGH  
NSNMGSDSGTMSDSEENVAARASPEPELQLRPYQMEVAQPALEGNIIICLPTGSG  
KTRVAVYIAKDHLDKKKASEPGRVIVLVNKLVLVEQLFRKEQPFLLKMYRVIGLSG  
DTOLKISFEVVKSCDIIISTAQILNLSLNLNENGEDAGVQSDSEFIIDECHITNK  
EAVYNNIMRHYLMOKLKNRLKKNKVPVLPQILGLTASPGVGAATKQAKAEHILK  
LCANLDAFTIKTVKENLDQLKNQIQEPCKFADATREDPFKEKLEIMTRIQTYYCQ  
MSPMSDFGTQPYEQWAIQMEKKAAGNRKESVCAEHLRKTNAALQINDTIRMIDAYT  
HLETFYNEEKDKKFAVIEDSDGDEYCDGDEDEDLKKPLKDETDRFLMTLFFE  
NNKMLKRLAENPEYENEXLTKLRNTIMEQYTRTEESARGIIFTKRQSAVALSQWITE  
NEKFAEVGVAAHLLIGAGHSEFEKPTQNEQKEVISKFRGKINLLIATVAEBGLDI  
KECNIVIRYGLVTNEIAMVOARGRADESTYVLVAHSGSGVIERETVNDFREKMYK  
AHCVOQMKREBYAHKILELOMQSIMEKMKTKRNIAKHVKNPSLITFLCKNSVLA  
CSGEDIHVIEKMHVNMTPPEFKELYVRENKALQKKCADYOINGEIIICKGOAWGTMM  
VHKGLDPLCLKIRNFVVVPKNNSTYKQYKKWVELPITFPNLDSBCCLPFSDDE"  
1148..1201  
/note="Region: DEAD-like helicase domain"  
misc\_feature 1481..1627  
/note="Region: DEAD/DEAH box"  
misc\_feature 2606..2647  
/note="conserved C-terminal helicase region; Region:  
helicase C"  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.64e-311 length: 3373  
Score: 5285.00 Matches: 1020  
Percent Similarity: 99.7% Conservative: 2  
Best Local Similarity: 99.5% Mismatches: 3  
Query Match: 99.5% Indels: 0  
DB: 8 Gaps: 0  
US-09-515-363C-2 (1-1025) x AY017378 (1-3373)  
QY 1 MetSerAsnGIyTyrSerThrAspGIuAsnPheArgTyrLeuIleSerCysPheArgAla 20  
| | | | |  
Db 155 ATGTCGAATGGGTATTCACAGACGAGAATTTCCTATCTCATCTCGTTCAGGGCC 214  
QY 21 ArgValLysMetTyrIleGIuValGIuProValLeuAspTyrLeuThrPheLeuProAla 40  
| | | | |  
Db 215 AGGTGAATAATGTACATCATCGAGCTGAGCCTGTGGACTACCTGACCTTCTGCCTGCA 274  
QY 41 GIuValLysGIuLysIleGIuArgThrValAlaThrSerGIyAsnMetGIuAlaValGIu 60  
| | | | |  
Db 275 GAGGTGAAGGAGCAGATTCAAGAGCAGACTCGCCACCTCCGGAAACATGCAGGAGTTGAA 334  
QY 61 LeuLeuLeuSerThrLeuGIuLysGIyValTrrpHisLeuGIyTrrpThrArgGIuPheVal 80  
| | | | |  
Db 335 CTGCTGCTGAGCACCTTGAGAAAGGAGTCTGGCACCCTTGCTTGACTCGGGAATTCGTG 394  
QY 81 GIuAlaLeuArgArgThrGIySerProLeuAlaAlaArgTyrMetAsnProGIuLeuThr 100  
| | | | |  
Db 395 GAGGCCCTCCGAGAACCGGACCCCTCTGCGCCGCTACATGAACCTGAGCTCACG 454  
QY 101 AspLeuProSerProSerPheGIuAsnAlaHisAspGIuTyrLeuGIuLeuLeuAsnLeu 120  
| | | | |  
Db 455 GACTGCCCTCTCCATCGTTTGAAGACGTCATGATGATATCTCCAACCTGCTGAACCTC 514  
QY 121 LeuGIuProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGIu 140  
| | | | |  
Db 515 CTTGAGCCCACTCTGTGTGACAAAGCTTCTAGTTAGAGACGTTCTTGATAGTGCAATGGAG 574  
QY 141 GIuGIuLeuLeuThrIleGIuAspArgAsnArgIleAlaAlaGIuAsnAsnGIyAsn 160  
| | | | |  
Db 575 GAGGAACGTGTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAATGGAAT 634  
QY 161 GluSerGIyValArgGIuLeuLeuLysArgIleValGIuLysGIuAsnTrrpPheSerAla 180  
| | | | |  
Db 635 GAATCAGGTGTAGAGAGACTATAAAAGGATTGTGCAGAAAGAAACTGTTCTCTGCA 694  
QY 181 PheLeuAsnValLeuArgGIuThrGIyAsnAsnGIuLeuValGIuGIuLeuThrGIySer 200

|||||  
Db 695 TTCTGAATGTTCTTCGCAAAACAGAAACAATGAATCTGTCCAAGAGTTAACAGGCTCT 754  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 755 GATTGCTCAGAAAGCAATGCAGAGATTGAGAATTATCACAAAGTTGATGGTCTCAAGTG 814  
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240  
Db 815 GAAGAGCAACTCTTTCAACCAACAGTTCAGCCAAATCTGGAGAAAGAGGTCTGGGGCATG 874  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
Db 875 GAGATAAATCATCAGAAATCATCTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACA 934  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 935 AGTTTGGCAGAAAGAGTGTCACTGCTTAGATGAAGTCTTGACATTAACAGCAACATG 994  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
Db 995 GGCAGTGATTCAGGCACCACTGGGAAGTGATTCAGATGAAGAGATGTGGCAGCAAGACA 1054  
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
Db 1055 TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAAGTTGCCCGCAGCCCTTG 1114  
QY 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340  
Db 1115 GAAGGGAAGATATCATCATCTGCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTT 1174  
QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysValAsnGluProGlyLysValIle 360  
Db 1175 TACATTTGCCAAGGATCATTAGACAAGAAAAAGCATCTGAGCCTGAAAAAGTTATA 1234  
QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380  
Db 1235 GTTCTTGTCATTAAGTACTGTAGTTGAACAGCTCTCCGCAAGAGTCCCAACCATTT 1294  
QY 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400  
Db 1295 TTGAAGAATGATCGTATCGTGTATTGGATTAAAGTGATACCCAACTGAATAATCATTT 1354  
QY 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
Db 1355 CCAGAAAGTTGTCAAAGTCCGTGTATATTATATCAGTACAGCTCAAACTCTTGAAAACTCC 1414  
QY 421 LeuLeuAsnLeuGluAsnGlyLysAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
Db 1415 CTCTTAAACTTGAAAAATGGAGAAGATGCTGTGTTCATTTGTCAAGCTTTTCCTTCATT 1474  
QY 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460  
Db 1475 ATCATTTGATGAATGTCATCACAACCAAGAGCAGTGTATAATAACATCATGAGGCAT 1534  
QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480  
Db 1535 TATTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATCCC 1594  
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500  
Db 1595 CTTCTCAAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAGCAAGCC 1654  
QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
Db 1655 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACT 1714  
QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540  
Db 1715 GTTAAAGAAAACTTGATCAACTGAAAAAACCAATAACAGAGCCATGCAAGAGTTGCC 1774  
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560  
|||||

Db 1775 ATTGCAGATGCACCAAGAGATCCATTTAAGAGAAACTTCTAGAAATAATGACAAAG 1834  
QY 561 IleGlnThrTyrCysGlnMetSerPrometSerAspPheGlyThrGlnProTyrGluGln 580  
Db 1835 ATTCAAACTTATTGTCCAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGAACAA 1894  
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
Db 1895 TGGGCCATTCAATGGAAAAAAAGCTGCAGAAAAGAGAAATCCGAAAGAAAGTGTGTGT 1954  
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
Db 1955 GCAGAACATTGAGGAAGTACAAATTAAGGCCCTTACAAATTAAAGACAAATTCGAATGATA 2014  
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640  
Db 2015 GATGCCGATACTCATCTTGAAACTTCTATTAATGAAGAGAAAGATTAAGAGTTTGCACTC 2074  
QY 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 2075 ATAGAAAGATGATAGTATGAGGGTGGTGATGATGAGTATTGTGATGGTATGAAGATGAG 2134  
QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2135 GATGATTTAAAGAAACCTTTGAAACTCGATGATAACAGATAGATTCTCATGACTTATT 2194  
QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2195 TTGAAAAACAATAAATGTTGAAAAAGCTGGCTGAAAAACCAAGATATGAAAAAG 2254  
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720  
Db 2255 CTGACCAATTTAAGAAATTAACCATTAATGAGCAATATACTAGAGCTAGGAATCAGACAGA 2314  
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740  
Db 2315 GGAATAATCTTTACAAAAACAGACAGAGTGCAATATGCCGTTCCAGTGGAATTACTGAA 2374  
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760  
Db 2375 AATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGGAGCTGCAGACAGC 2434  
QY 761 SerGluPheLysPrometThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2435 AGTGAAGTCAAAACCATGACACAGAAATGAACAAAAAGATCATTAGTAAATTTCCGACT 2494  
QY 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2495 GGAAAAATAAATCTGTTATCGCTACACAGTGGCAGAAAGGTCTGATATTAAAGAA 2554  
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2555 TGTAACATTTGTTATCCGTTATGGTCTCGTCAACCAATGAATAAGCCATGGTCCAGCCCGT 2614  
QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2615 GGTCAAGCCAGAGCTGATGAGACACCTACGTCCTGCTGCTCACAGTGGTTCAAGAGTT 2674  
QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860  
Db 2675 ATCGAAGTGAGACAGTTAATGATTTCGAGAGAAAGATGATGATATAAGCTATAATTGT 2734  
QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2735 GTTCAAAATATGAAAACCAAGAGATATGCTCATAGATTTTGAATTACAGATGCAAAAGT 2794  
QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900  
Db 2795 ATAATGGAAGAAAGAAATGAAAACCAAGAGAAATATGCAAGCATTAACAGAAATCAACA 2854  
QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920  
Db 2855 TCACATAATACTTCTTCGCAAAAACTGCAGTGTGCTAGCCTGTCTGGGGAAGATATC 2914

Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
Db 2915 CATGTAATTGAGAAATGCAATCACTCATATGACCCAGAAATTCAAGGAAGTTTACATT 2974  
Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 2975 GTAAGAGAAAACAAGCACTGCAAAAGAGTGTGCCACTATCAATAATAGTGAAATC 3034  
Qy 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3035 ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGGCCT 3094  
Qy 981 CysLeuLysIleArgAsnPheValValAlaPheLysAsnAsnSerThrLysLysGlnTyr 1000  
Db 3095 TGTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCAACAAAGAAACAATAC 3154  
Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020  
Db 3155 AAAAAGTGGGTAGAATTACCTATCACTTTCCCAATCTTGACTATTCAAGATGCTGTTTA 3214  
Qy 1021 PheSerAspGluAsp 1025  
Db 3215 TTTAGTGATGAGGAT 3229  
RESULT 11  
AY075132 3078 bp mRNA linear ROD 20-MAY-2002  
LOCUS Mus musculus HELICARD mRNA, complete cds.  
DEFINITION AY075132  
ACCESSION AY075132 GI:18698980  
VERSION AY075132.1  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 3078)  
AUTHORS Kovacsovics,M., Martinon,F., Micheau,O., Bodmer,J.L., Hofmann,K.  
and Tschopp,J.  
TITLE Overexpression of Helicard, a CARD-Containing Helicase Cleaved during Apoptosis, Accelerates DNA Degradation  
JOURNAL Curr. Biol. 12 (10), 838-843 (2002)  
PUBMED 12015121  
REFERENCE 2 (bases 1 to 3078)  
AUTHORS Kovacsovics,M., Hofmann,K. and Tschopp,J.  
TITLE Direct Submission  
JOURNAL Submitted (24-JAN-2002) Biochemistry, University of Lausanne, ch  
des Boveresses 155, Epalinges, VD 1066, Switzerland  
FEATURES  
source location/Qualifiers  
1..3078  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
1..3078  
/note="helicase; contains two CARD domains"  
/codon\_start=1  
/product="HELICARD"  
/protein\_id="AAL77205.1"  
/db\_xref="GI:18698981"  
/translation="MSIVCSAEDSFRNLILFRPRLMKYIQVEPVLDHLIFLSAETKE  
QILKINTCGNTSAEELLSTLEQGWPILGQMFVEALBHSGNPLAARYVKPTLTDL  
PSPSSETAHDECLHLTLTLPDLVDKLINDVLTCEBGLTVEDNRIRISAAGNSGN  
ESGVELLRIRIVQKENVSTFLDVLROTDNALFQELTGGCPEDNTDLANSSHRDGP  
AANECLLPAVDESSLETEAMVNDILPEASCTDSSVTTESDTSIAEGSVSCFDESLGH  
NSNMGDSGTMSDSDESIVITKRVSPPELQLRPYQMEVAQPALDGNIIICLPTGS  
GKTRVAVYITKDLHDKKQASESGKIVLVNRYMLAEQLFRKEFNPLYLKKYRIIGLS  
GDTQLKISPEVVKSYDVIISTAQILENSLNLBESGDDDGVLSDFSLIIDECHHTN  
KEAVYNNIMRRLKOKLRNNDLKKONKPAIPLOILGITASPGVGAACKQSEAEKHIL  
NICANLDAFTIKTVKENLQGLKHQIKEPCKFVIADDTRENPFKEKLEIMASIQTYC  
OKSPMSDFGTOHYEOWAIOMEKKAADGNRKDYCAEHLRKYNALQINDTIRMIDAY  
SHLETFTDEKEKKFAVLNDSKSDDEASSCNDQLKGDKVKSILKDETDFLMLNLFED  
NKKMLKKLAENPKYENEKLIKLRLNTILAEQFTSSESSRGIIFTKROSTYVALSQWIME

NAKFAEVGVAHHLIGAGHSSEVKEMTQTQKEVISKFRGTGINLLIATTVAEBGLDI  
KECNIVIRYGLVTNEIAMVQARGRADESTYVLVTSSGSGYTEREIVNDFREKMYK  
AINRVONMKPEBYAHNILELOVQSLIEKKMKVRSIAQYNDNPSLITLCKNCSMLV  
CSGENIHVIEKMHVNMTPFEKGLYVRENKALQKKFADYQTNGEIICKCGAMGTMM  
VHKGLDPLCLKIRNFVNPNKNSPKQYKKWVLPPIRFPDLDYSEYCLYSDDED"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.78e-245 Length: 3078  
Score: 4199.00 Matches: 818  
Percent Similarity: 87.0% Conservative: 75  
Best Local Similarity: 79.7% Mismatches: 131  
Query Match: 79.1% Indels: 2  
DB: 9 Gaps: 2

US-09-515-363C-2 (1-1025) x AY075132 (1-3078)

Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
Db 1 ATGTGATTTGTTCTGTCTGACAGAGACAGCTTCAGGAATCTCATTTCTTACAGGCC 60  
Qy 21 ArgValIysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
Db 61 AGGCTGAAATGTATCATTCAGGTGAGCCAGTCTGACACCACCTCATCTTCTGTGCA 120  
Qy 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 121 GAAACCAAGACAGACATCTTAAAGATCAACACCTGTGTGTAACACCAAGCCGCGAGAA 180  
Qy 61 LeuLeuLeuSerThrLeuGlnLysGlyValTrpHisLeuGlyTyrTrpThrArgGluPheVal 80  
Db 181 CTGCTGTGAGACACCTTGAGACAGGACAATGGCCTCTGGAGTAGACGACGATGTTGCTG 240  
Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
Db 241 GAGGCCCTAGACACAGTGGCAATCCCTAGCCGCGCTATGTCAAAACCACACTCACT 300  
Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 301 GATCTGCCCTCTCTCTCTGAGACTGCCCATGACGAGTGTCTCCACTTGTGACCCCTC 360  
Qy 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140  
Db 361 CTCACGCCCACTTTGGTGACAAACTTGTGATTAACGATGCTTGACACTGCTTCGAG 420  
Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
Db 421 AAGGACTATTGACAGTGCAGACAGAAATCGATTCTGCTGCAGAAACAGCGGAAT 480  
Qy 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180  
Db 481 GAGTCAGGTGTAAGAGAGCTGCTGAGAAGATGTGCAGAGAGAAAAGTGTCTTCTACC 540  
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 541 TTCTGGATGTTCTGCGCCAAACTGAAATGATGCACTATTCCAAGAACTTAACAGGTGA 600  
Qy 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 601 GGCTGCCCAGAGACAAACACAGACTGGCTAATCTGCTCACAGAGATGGCCTGCAGCT 660  
Qy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240  
Db 661 AATGAGTGTCTTCTGCTGCTGTGATGAGTCAAGTCTTGAGACAGAGGCTGGAACGTA 720  
Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValIleSerGluSerAspThr 260  
Db 721 GACGACATATTACAGAGGCTTCTGTACAGATTGCTGTGACCACAGAAATCAGACACA 780  
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 781 AGTTTGCAGAGAGAGAGTGTCACTGCTTGCATGAAGTCTTGACATTAACAGCAACATG 840

QY	281	GLYSERASPSERGLYTHRMETGLYSERASPSERASPGIULASUNVAL---ALALAARG	299
DB	841	GGCAGGATTCAGGACCATGGGAAGTGAATTCAGATGAAGTGTCTATCCAGACAAAAAGA	900
QY	300	ALASERPROGLUPROGLULEUGLNULEUARGPROTYRGIMETGLUVALAGINPROALA	319
DB	901	GTATCCCCCGAGCCAGAACTGCAGCTCAGGCCCTTACCAATGGAAGTGCCCAACCACT	960
QY	320	LEUGLGLYLYSASNILEILEILECYLSLEUPROTHRGLYSERGLYLSRTHRARGVALA	339
DB	961	CTAGATGGGAGAATAATTATATCTGCCTCCCCACGGGAGTGGGAAAAACCAAGTGGCT	1020
QY	340	VALTYRILEALALYSAPHISLEUASPLYSLSYLSVALASERGLUPROGLYLYSVAL	359
DB	1021	GTTTACATCACCAAGATCACTTAGACAAAGAAGCAGGCATCTGAATCCGGGAGGTT	1080
QY	360	ILEVALLEUVALASNLVSVALLEUVALGLUGLNULEUPHEARGLYSGLUPHEGINPRO	379
DB	1081	ATCGTCTTGTCAAATAGGTAATGTTAGCAGAACAACTTTCGAAAAAGATTCAACCA	1140
QY	380	PHLEULYLSYSTRPYRRARGVALILEGLYLEUSERGLYASPTHRGILNULEUYILESER	399
DB	1141	TATTTGAAGAATGTTATCGAATTATTGGATTAAAGTGGCATACCAGCTGAATAATCA	1200
QY	400	PHPROGLUVALVALYSERCYASPILEILEISERTHRALAGINILEUGLUSN	419
DB	1201	TTTCAGAAAGTTGICAAATCTTACGATGTTATTATCAGCACTGCTCAAAATCCTTGAAAC	1260
QY	420	SERLEULASNLUGLUSNGLYGLIUSPALAGLYVALGINLEUSERASPSERLEU	439
DB	1261	TCCCTCTTAATCTGGAGAGTGAGACGATGACGGTGTGCAGCTGTCAAGACTTCTCTC	1320
QY	440	ILEILEILEASPGIUCYSHISHISTHRASNLYSGLUVALTYRRASNANILEMETARG	459
DB	1321	ATTATCATTTGATGAGTGCCATCACACCAAGAAGGAGGTCTATAACAACATCATGAGA	1380
QY	460	HISTYRLEUMETGLINYSLEULYSASNANARGLEULYSGLUSNLSPROVALILE	479
DB	1381	CGATATTTGAAGCAGAACTGAGAAACAATGACCTCAAGAAAAACAACCAACCGCAATT	1440
QY	480	PROLEUPROGINILEUGLYLEUTHRALASERPROGLYVALGLYALATHRLYSGIN	499
DB	1441	CCCCTGCCGACAGATACTAGCACTGACAGCTTCACTGGTGTGGAGCAGCCAAAAAGCAG	1500
QY	500	ALALYSALAGLUHISILEULEULYSLEUCYSALASNLUASPALAPHETHRIELYS	519
DB	1501	TCTGAGGCTGAATAACATATTTAAATATATATGTGCCATCTTGATGCCCTTACCATTA	1560
QY	520	THRVALYSGIUSNLUASPNLEUASPNLEULYSASNGINILEGINLUPROCYLSYLSYPHE	539
DB	1561	ACAGTGAAGAAGAACTTGGTCAACTCAAAACACCAATTAAGGAACCATGCAGAGAAATTT	1620
QY	540	ALALILEALASPALATHRARGLUASPPROPHELYSGLULYSLEULEUGLULEMETTHR	559
DB	1621	GTGATTGCTGATGACACACGAGAAAAATCCATTAAAGAAAACTTCTAGAAATTATGCA	1680
QY	560	ARGILEGLINTHRTRYCYSGIMETSERPROMETSERASPBHEGLYTHRGINPROTYRGIU	579
DB	1681	AGCATTCAGACTTACTGCCAAAAAGTCCAAATGTCAAGTTTGGAAACCCCAACATTATGAG	1740
QY	580	GLINTRPALALEGLINMETGLULYSALALALALYSYSGIYASNARGLYSGIUNARGVAL	599
DB	1741	CAGTGGCCCATTCAAATGAGAAAAAGCTGCTAAAGACGGAATTCGCAAGATCGCGTC	1800
QY	600	CYSALAGLUHISLEUARGLYSTRASNGIUALALEUGLINLEASNASPTHRILEARGMET	619
DB	1801	TGTGACGAACATTTGAGGAAGTAAACAAGAGCCCTACAAATCAACGACACGATCCGAATG	1860
QY	620	ILEASPALATYRTHRHSILEUGLUTHRPHETRYASNGLUGIULYSASPLYSLSYPHEALA	639
DB	1861	ATTGATGCATATAGCCACCTGGAGACATTTCTACACTGATGAGAAAAAGAAAGATTGCA	1920
QY	640	VALILEGLUASPSERASPGIULGLIYASPSASPGIUTYRCYASASPGIYASPGIUNASP	659

Db	1921	GTCTC--AATGACAGCGACAGAGTGAAGCGCCAGCAGTTGCAATGACCAACTT	1977
Qy	660	GLuAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu	679
Db	1978	AAGGCGCATGTAAAGAAATCTTTGAACCTGCAGAAACGATGAATTTCTCATGAATTTG	2037
Qy	680	PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu	699
Db	2038	TTCTTTGATTAACAAGAAATGTTGAAAAAACTAGCTGAAAAACCAAAATACGAGATGAA	2097
Qy	700	LysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluSerAla	719
Db	2098	AAACTCATTAATTAAGAAACACGATACTGGAACAATTCACAAAGGTCTGAGGAGTCTCC	2157
Qy	720	ArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThr	739
Db	2158	CGAGGAATTAATTTTACAAAAACAGACAGACACTTACGCACTTCCAGTGCATCATG	2217
Qy	740	GluAsnGluLysPheAlaGluValGlyValLysAlaHisIleLeuIleGlyValaGlyHis	759
Db	2218	GAAATGCAAGTTTGCGGAAGTTGGAGTCAAGCCGATCACCTGATTGGCGCGGCGAC	2277
Qy	760	SerSerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArg	779
Db	2278	AGCAGTGAAGTCAAGCCCATGACTCAGACTGAACAAAAAGAGTCATTAGTAAATTTCCG	2337
Qy	780	ThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys	799
Db	2338	ACTGGCGAAATAAATCTGCTTATCGCTACGACGGTGGCAGAGGAAGCCTGATATCAAA	2397
Qy	800	GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla	819
Db	2398	GAGTGCATATGTTGTTATTCGTTATGCGCTTGTCACGAACGAGATAGCCATGGTCCAGGCC	2457
Qy	820	ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly	839
Db	2458	CGGGGTGAGCCAGAGCTGATGMAAGCAGCTATGCTGTCACCAGCAGTGGCTCAGGA	2517
Qy	840	ValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis	859
Db	2518	GTTACCGAACGGAGATTGTTAATGATTTCCGAGAGAAAGATGATGATTAAGCTATTAAAC	2577
Qy	860	CysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGln	879
Db	2578	CGTGTCAAAACATGAACACAGAGAGTATGCACATAAGATTTTGAATTGCAGGTGCAA	2637
Qy	880	SerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsn	899
Db	2638	AGTATCCTGMAAAGAAATGMAAGTCAAAAGACATTGCCAAAGCAATACACAGCAAT	2697
Qy	900	ProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAsp	919
Db	2698	CCATCGTTAATACACTTCTGCAAAAAATTGTGACATGCTGCTGCTCGGAGAAAC	2757
Qy	920	IleHisValIleGluLysMetHisIleValAsnMetThrProGluPheLysGluLeuTyr	939
Db	2758	ATCCATGTCATTGAGAAGATGCATCATGTCAATATGACACCAAGATTCAAGGACTCAAC	2817
Qy	940	IleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlu	959
Db	2818	ATTGTAAGAGAAAACAAAAGCACTGCMAAAGAAATTTGCTGATTATCAGACCAATGAGAG	2877
Qy	960	IleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeu	979
Db	2878	ATTATCTGCAAGTGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGTTTAGATTGG	2937
Qy	980	ProCysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGln	999
Db	2938	CCTTGCTTAATAAAGAATTTGTAGTCAATTTCAAAATAACTCACCGAAGAAACAG	2997
Qy	1000	TyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys	1019



Db 1795 ACAGTGAAGAAGATCTTGGTCAACTCAACACACCAATAAAGAACCATGCAAGAATTT 1854  
Qy 540 AlaIleAlaAspAlaThrArgGluAspProPheIysGluIysLeuLeuGluIleMetThr 559  
Db 1855 GTGATTGCTGATGACACACGAGAAAATCCATTAAAGAGAACTTCTAGAAATTATGCA 1914  
Qy 560 ArgIleGlnThrTyrCysGlnMetSerPrometSerAspPheGlyThrGlnProTyrGlu 579  
Db 1915 AGCATTCAGACTTACTGCCAAAAAGTCCAAATGTCAGATTGGAAACCCAAACATTATGAG 1974  
Qy 580 GlnTrpAlaIleGlnMetGluIysAlaAlaIysIysGlyAsnArgIysGluArgVal 599  
Db 1975 CAGTGGCCATTCAAAATGAGAAAAAGCTGCTAAAGACGGAATCGCAAAGATCGCGTC 2034  
Qy 600 CysAlaGluHisLeuArgIysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619  
Db 2035 TGTGCAGAACATTTGAGGAAGTACAAAGCCCTACAAATCAACGACACGATCCGAATG 2094  
Qy 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluIysAspIysPheAla 639  
Db 2095 ATTGATGATATAGCCACCTGAGGCATTTCTACACTGATGAGAAAAGAAAAGATTGCGA 2154  
Qy 640 ValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659  
Db 2155 GTTCCTC---AATGACAGCGACAGAAGTGAAGAGGCCACAGTTCGAATGACCAACTT 2211  
Qy 660 GluAspAspLeuIysIysProLeuIysLeuAspGluThrAspArgPheLeuMetThrLeu 679  
Db 2212 AAGGCGCATGTAAAGAAATCTTGAACCTGGACGAAACGATGAATTTCTCATGAATTTG 2271  
Qy 680 PhePheGluAsnAsnIysMetLeuIysArgLeuAlaGluAsnProGluTyrGluAsnGlu 699  
Db 2272 TTCTTGTGATAACAAGAAATGTTGAAAAAAGCTGAAAAACCAAAATACGAGAATGAA 2331  
Qy 700 LysLeuThrIysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGlnSerAla 719  
Db 2332 AAACTCATTAATAATGAAGAAACAGATACTGAAACAATTCAAGGTCTGAGAGTCTCTCC 2391  
Qy 720 ArgGlyIleIlePheThrIysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThr 739  
Db 2392 CGAGGAATTATTTTACAAAAACGACAGACAGACCTTCCAGTGGATCATG 2451  
Qy 740 GluAsnGluIysPheAlaGluValGlyValIysAlaHisIleuIleGlyAlaGlyHis 759  
Db 2452 GAAATGCAAGTTTTCGGAAGTTGAGTCAAAAGCCGATCACTGATTGGCCGCGGCAC 2511  
Qy 760 SerSerGluPheIysPrometThrGlnAsnGluGlnIysGluValIleSerIysPheArg 779  
Db 2512 AGCAGTGAAGTCAAGCCCATGACTCACTGAACAAAAAGAGTCATTAGTAATTTGCGC 2571  
Qy 780 ThrGlyIysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys 799  
Db 2572 ACTGGCGAAATAATCTGCTTATGCTACGACGGTGGCAGAGGAAGCCTGATATCAAA 2631  
Qy 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819  
Db 2632 GAGTGCATATTGTTATTCGTTATGGCCTTGTCAAGACAGATAGCCATGCTCCAGGCC 2691  
Qy 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839  
Db 2692 CGGGGTCGAGCCAGAGCTGATGAAGCAGTATGTCCTGGTCAACGACAGTGGCTCAGGA 2751  
Qy 840 ValIleGluHisGluThrValAsnAspPheArgGluIysMetMetTyrIysAlaIleHis 859  
Db 2752 GTTACCGAACGGAGATTGTTAATGATTTCCGAGAGAAAGATGATATAAAGCTATTAAAC 2811  
Qy 860 CysValGluAsnMetIysProGluGluTyrAlaHisIysIleLeuGluLeuGlnMetGln 879  
Db 2812 CGTGTCAAAACATGAAACAGAGAGTATGACATAAAGATTTTGAATTTGACAGTGCAC 2871  
Qy 880 SerIleMetGluIysIysMetIysThrIysArgAsnIleAlaIysHisTyrIysAsnAsn 899

Db 2872 AGTATCCTGGAAGAAAGAAATGAAGTCAAAAGAGCATTGCAAGCAATACACGACAAAT 2931  
Qy 900 ProSerLeuIleThrPheLeuCysIysAsnIysSerValLeuAlaCysSerGlyGluAsp 919  
Db 2932 CCATCGTTAATAACACTTCTCTGCAAAAAATTGTAGCATGCTGCTGCTCGGAGAAAAAC 2991  
Qy 920 IleHisValIleGluIysMetHisHisValAsnMetThrProGluPheIysGluLeuTyr 939  
Db 2992 ATCCATGTCATTGAGAGAGATGCATCATGTCTCATATATGACACCAAGAAATCAAGGACTTAC 3051  
Qy 940 IleValArgGluAsnIysAlaLeuGlnIysIysCysAlaAspTyrGlnIleAsnGlyGlu 959  
Db 3052 ATTGTAAGAGAAAAACAAGCACTGCMAAAGAAATTTGCTGATTATCAGACCAATGAGAG 3111  
Qy 960 IleIleCysIysCysGlyGlnAlaTyrGlyThrMetMetValHisIysGlyLeuAspLeu 979  
Db 3112 ATTATCTGCAAGTGTGGCCAGGCTTGGGAACAATGATGGTGACACAAGCTTTAGATTGG 3171  
Qy 980 ProCysLeuIysIleArgAsnPheValValPheIysAsnAsnSerThrIysIysGln 999  
Db 3172 CCTTGTCTTAAATAAGAAATTTGTAGTCATTTCAAAAATACTCACCGAAGAACAG 3231  
Qy 1000 TyrIysIysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019  
Db 3232 TACAAGAAAGTGGGTGAATTGCCCTATCAGATTTCCTGATCTTGACTACTCAGAATACTGC 3291  
Qy 1020 LeuPheSerAspGluAsp 1025  
Db 3292 TTGTATAGTGATGAAGAT 3309  
  
RESULT 13  
AF374384 3771 bp mRNA linear ROD 02-MAY-2002  
LOCUS AF374384  
DEFINITION Mus musculus strain C57BL/6J melanoma differentiation associated  
ACCESSION AF374384  
VERSION AF374384.1 GI:20385627  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
  
REFERENCE  
AUTHORS Kang,D.-C. and Fisher,P.B.  
TITLE Murine homolog of melanoma differentiation associated gene-5 - a  
JOURNAL DEXH group RNA dependent ATPase  
REFERENCE 1 (bases 1 to 3771)  
AUTHORS Kang,D.-C. and Fisher,P.B.  
TITLE Direct Submission  
JOURNAL Submitted (26-APR-2001) Urology, Columbia University, P & S, 630  
FEATURES  
source location/Qualifiers  
1. 3771  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
235..3312  
/note="RNA-dependent ATPase; predicted molecular weight:  
115954.50 Da"  
/codon\_start=1  
/product="melanoma differentiation associated gene 5-like  
protein"  
/protein\_id="AAM21359.1"  
/db\_xref="GI:20385628"  
/translation="MSIVCSAEDSFRLILFRRRLKMYIQVEPVLDHLIFLSAETKE  
QILKINTCGNTGSAEILLSTLEOGWPLGWTQMFVEALEHSGNPLAARYVKPTLTDL  
PSPSSETAHDECHLLTLQPTLVKLLINDVLDTCFEKGLTYEDRNRIISAAGNSGN  
ESGVRELRLRIIVOKENWESTFLDVLRGTGNDALFOELTGCGCPEDNTDLANSSHRDGP  
AANECLLPAVDESLETEAMVNDILPEASCTDSSTTESDTSLAEGSVSCFDESLGH  
NSNMGRDSGTMSDSDSVIQTKRVSPEPELQLRPYQMEVAQRPALDGKNIILCLPTGS

GKTRVAVYITKDHLDKKQASESGKIVLVNKMALBQLFRKENPYLKKWYRIIGLS  
GDTQKISFPEVVKSYDVIIISTAQILENSILNLESDDDGVOQLSDFSLIIDECHHTN  
KEAVYNNIMRRYLKOKLRNNDLKQKPAIPLOILGLTASPGVGAARKQSEAEKHIL  
NICANLDAFTIKTYKENLGOQLKHQIKPECKKFVIAADTRENPFKEKLEIMASIQTVC  
OKSPMSDFGTOHYEQWAIQMEKKAADGNRQVCAEHLRKYNREALQINDTIRMIDAY  
SHLEAFYDEKEKKFAVLNDSKSDDEASSCNDQLKGDVKSLKLDETDEFLMNFED  
NKKMLKKLAENPKYENMKLIKLRNTILBQFTRSEESSRGIIFTKRQSTYALSQWIME  
NAKFAEVGVKAHHLIGAGHVSSEKVPMTQTEQKEVISKFRTEINELIATTVAEBGDI  
KECNIVIRYGLVTNETIAMQVARGARADESTYVLVTSSSGVTEREIVNDFREKMMYK  
AINRVQNMKPEEYAHKILEQVQSILEKMKYKRSIAKQYNDNPSTLLCKNCSMLV  
CSGENIHVIEKMHVNMTPPEFKGLYIVRENKALQKKFADYQTNGEIICKGQAWGTMM  
VHKGLDLPCLKIRNFVNFKNNSPKKQYKKWELPIRFPDLDYSEYCLYSDDED"  
misc\_feature  
1228..1245  
/note="ATP binding motif"  
misc\_feature  
1564..1575  
/note="DEXH helix signature; unclassified site"  
polya\_signal  
3755..3760  
ORIGIN

Alignment Scores:  
Pred. No.: 9.39e-245 Length: 3771  
Score: 4194.00 Matches: 817  
Percent Similarity: 86.9% Conservative: 75  
Best Local Similarity: 79.6% Mismatches: 132  
Query Match: 79.0% Indels: 2  
DB: 9 Gaps: 2

US-09-515-363C-2 (1-1025) x AF374384 (1-3771)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
DB 235 ATGTCGATTGTCGTCGACAGACAGCTTCAGGAATCTCATCTTATTCTTCAGGCC 294  
QY 21 ArgValIleMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
DB 295 AGGCTGAATAATGATCATTCAGGTGAGCCAGTGCTGGACCACCTCATCTTCTGCTGCA 354  
QY 41 GluValIleGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
DB 355 GAAACCAAGACAGATTCTTAATAAAGATCAACACCTGTGTAAACACCAGCGCGAGAA 414  
QY 61 LeuLeuSerThrLeuGluIleGlyValTyrPheIleGlyTyrThrArgGluPheVal 80  
DB 415 CTGCTGTCAGCACCTTGAGCAGGACAATGGCCTCGGAGATGAGCAGAGATGTCGTG 474  
QY 81 GluAlaIleuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
DB 475 GAGGCCCTAGACACAGATGGCAATCCCTAGCGCGCTATGTCAAAACCCACACTCACT 534  
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
DB 535 GATCTGCCCTCTCCTCTCTGAGACTGCCCATGACGAGTGTCTCCACTTGCTGACCCCTC 594  
QY 121 LeuGlnProThrLeuValAspIleLeuLeuValArgAspValLeuAspIleCysMetGlu 140  
DB 595 CTCACGCCCACTTGTGTGACAAACTTCTGATTAAAGATGTCTTGACACTTGTCTTCGAG 654  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaAlaGluAsnAsnGlyAsn 160  
DB 655 AAGGACTAATTGACAGCTGCAAGACAGAAATCGATTTCGTGACGAAACAGCGGGAAT 714  
QY 161 GluSerGlyValArgGluLeuLeuIleAspArgIleValGlnIleGlyAsnTyrPheSerAla 180  
DB 715 GAGTCAGGTGAAGAGAGCTGCTGAGAAGATGTGCAAGAAAGAAACTGTTTCTACC 774  
QY 181 PheLeuAsnValIleuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
DB 775 TTCCTGATGTTCTGCGCCAAACTGGAATATGATGACTATTCCAAGACTAACAGGTGGA 834  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
DB 835 GGCTGCCCAAGACACACAGACTTGCTTAACCTGCTTCACAGAGATGGGCTGCGAGCT 894

QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIleGlyValTyrGlyMet 240  
DB 895 AATGATGTCTTCTGCTGCTGTCGATGATGATCAAGTCTGAGACAGAGGCTGGAACGTA 954  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValIleSerGluSerAspThr 260  
DB 955 GACGACATATTACACAGAGCTTCTGTATACAGATTGCTGTGACACACAGATCAGACACA 1014  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
DB 1015 AGTTGGCAGAGAAGATGTCAGCTGCTTCGATGAAGAAGCTTGACATACAGACAACATG 1074  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluAsnVal---AlaAlaArg 299  
DB 1075 GGCAGGGAATTCAGGCACCATGGGAAGTGATTGAGATGAAGAAGTGTCATCCAGACAAAGA 1134  
QY 300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319  
DB 1135 GTATCCCCGAGCCAGACTGCAAGCTCAGGCTTACCAATGGAAGTGGCCCAACCAAGCT 1194  
QY 320 LeuGluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAla 339  
DB 1195 CTGATGGGAAGATATTATTATCTGCTCCCAAGGGAGTGGAAACAGAGTGCT 1254  
QY 340 ValTyrIleAlaIleAspHisIleuAspIleLysIleLysAlaSerGluProGlyLysVal 359  
DB 1255 GTTTACATCACCAAGATCACTTAGACAAGAAGAACAGCATCTGAATCCGGGAAGTT 1314  
QY 360 IleValLeuValAsnIleValLeuLeuValGluGlnLeuPheArgIleGluPheGlnPro 379  
DB 1315 ATCGTTCTGTCAATAAGTAATGTTAGCAGAACAACTTTCGAAAGAATTCAACCCA 1374  
QY 380 PheLeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSer 399  
DB 1375 TATTGAAGAAATGTATCGAATTATTGAGTTAAGTGGGATACCAGCTGAATAATATCA 1434  
QY 400 PheProGluValValIleSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsn 419  
DB 1435 TTTCAGAAAGTTGTCAAACTTACGATGTATTATCAGCACTGCTCAAAATCTTGAAAC 1494  
QY 420 SerLeuLeuAsnLeuGluAsnGlyLysAlaAspAlaGlyValGlnLeuSerAspPheSerLeu 439  
DB 1495 TCCCTCTTAATCTGAGAGTGGAGACGATGACGGTGTGACGCTGTCAAGCTTCTCTC 1554  
QY 440 IleIleIleAspGluCysHisHisThrAsnIleGluAlaValTyrAsnAsnIleMetArg 459  
DB 1555 ATTATCATTTGATGATGATGCCATCACACCAAGAGAGGAGTCTATTAACAACATCATGAGA 1614  
QY 460 HisTyrLeuMetGlnIleLysLeuLysAsnAsnArgLeuLysIleGluAsnLysProValIle 479  
DB 1615 CGATTTTGAAGCAGAAAGCTGAGAAACATGACTCAAGAAACAAACCAACCAAGCCATT 1674  
QY 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGln 499  
DB 1675 CCCCTGCCGACAGATACAGACTGACAGCTTCACTGTGTGAGCAGCAAAAAAGCAG 1734  
QY 500 AlaIleAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519  
DB 1735 TCTGAGGCTGAATAACATATTTAAATATATGTGCCAATCTTGATGCCTTTAAACATTAAA 1794  
QY 520 ThrValIleGluAsnLeuAspGlnLeuLysAsnGlnIleGlnIleProCysLysLysPhe 539  
DB 1795 ACAGTGAAGAAGATCTTGTCACACTCAAAACACCAATAAAGAACCATGCAAGAAATTT 1854  
QY 540 AlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGlnIleMetThr 559  
DB 1855 GTGATTGCTGATGACACCCAGAGAAATCCATTTAAGAGAACTTCTGAATTAATGCGCA 1914  
QY 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579  
DB 1915 AGCATTCAGACTTACTGCCAAAAAGTCCAAATGTCAGATTTTGAAACCCCAACATTATGAG 1974  
QY 580 GlnTyrAlaIleGlnMetGluLysLysAlaIleLysLysGlyAsnArgLysGluArgVal 599

Db 1975 CAGTGGCCATTCAATGAGAAAAAGCTGCTAAAGACGGAATCGCAAGATCGCGTC 2034  
Qy 600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619  
Db 2035 TGTCGACAACATTGAGGAAGTACAACGAGCCCTACAAATCAACGACGATCCGAATG 2094  
Qy 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAla 639  
Db 2095 ATTGATGATATAGCCACCTGGAGGCATTCTACACTGATGAGAAAGAAAGAAAGTTCCGA 2154  
Qy 640 ValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659  
Db 2155 GTCTCTC--AATGACAGCGACAAGTGAATGATGACGAGCCAGTTCGAATGACCAACTT 2211  
Qy 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679  
Db 2212 AAGGCGATGTAAAGAAATCTTGAAAACCTGACGAAACGATGAATTTCTCAATGAATTTG 2271  
Qy 680 PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu 699  
Db 2272 TTCTTGATATAACAAGAAAATGTTGAAAAAACTAGCTGAAAAACCAAAATACGAAATGAA 2331  
Qy 700 LysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAla 719  
Db 2332 AAATCATTAATTAAGAAACGAGTACTGGAACAATTCAACAAGTCTGAGAGTCTCTCC 2391  
Qy 720 ArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThr 739  
Db 2392 CGAGGAATTATTTTACAAAAACACGACAGAGCACCTACGCACTTCCCGATGATCATG 2451  
Qy 740 GluAsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHis 759  
Db 2452 GAAATGCAAAAGTTTGGAGAGTTGAGAGTCAAAGCGCATCACTGATTGGCGGGGCGAC 2511  
Qy 760 SerSerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArg 779  
Db 2512 AGCAGTGAAGTCAGCCCATGACTCAGACTGAACAAAAGAAAGTCATTAGTAATTTCCG 2571  
Qy 780 ThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluLysAspIleLys 799  
Db 2572 ACTGGCGAAATTAATCTGCTTATCGCTACGACGGTGGCAGAGGAGGCTGATATCAA 2631  
Qy 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819  
Db 2632 GAGTGCATATTGTTATTCGTTATGCGCTTGTACGAAACGAGATAGCCATGGTCCAGGCC 2691  
Qy 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839  
Db 2692 CGGGGTGAGCCAGAGCTGATGAAAGCAGTATGTCGTGTCACGACGAGTGGCTCAGGA 2751  
Qy 840 ValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859  
Db 2752 GTTACCGAACGGAGATTGTTATGATTTCCGAGAGAGATGATGTATAAGCTATTAAAC 2811  
Qy 860 CysValGluAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGln 879  
Db 2812 CGTGTCAAAACATGAACACGAGAGTATGCACATAGATTTTGAATTGCAAGTGCAC 2871  
Qy 880 SerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsn 899  
Db 2872 AGTATCTCGAAAAAGAAATGAAGTCAAAAAGAGCATTCGCAAAAGCAATACACGACAAT 2931  
Qy 900 ProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyLysAsp 919  
Db 2932 CCATCGTTAATAACACTTCTCTGCAAAAATTGTAGCATGCTGTGCTCGGAGAAAAAC 2991  
Qy 920 IleHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyr 939  
Db 2992 ATCCATGTCATTGAGAAGATGCATCATATATGACACCAAGAAATTCAGAGGACTCTAC 3051  
Qy 940 IleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlu 959

Db 3052 ATTGTAAGAGAAAAACAAGCACTGCAAAAAGAAATTGCTGATTATCAGACCAATGGAGAG 3111  
Qy 960 IleIleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeu 979  
Db 3112 ATTATCTGCAAGTGTGGCCAGGCTTGGGGACAATGATGTGTGCACAAAGTTTAGATTGG 3171  
Qy 980 ProCysLeuLysIleArgAsnPheValValIlePheLysAsnAsnSerThrLysLysGln 999  
Db 3172 CTTGTCTTAAATTAAGAAATTTGTAGTCAATTTCAAAAATTAATCAACGAAAGAACAG 3231  
Qy 1000 TyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019  
Db 3232 TACAAGAAAGTGGTGAATTGCCCTATCAGATTCTCTGATCTTGACTACTCAGAATACTGC 3291  
Qy 1020 LeuPheSerAspGluAsp 1025  
Db 3292 TTGTATAGTGAAGAAT 3309  
RESULT 14  
BC080200  
LOCUS  
DEFINITION  
MUS musculus interferon induced with helicase c domain 1, mRNA  
(CDNA clone MGC:90959 IMAGE:3597815), complete cds.  
ACCESSION  
BC080200  
VERSION  
BC080200.1 GI:51329783  
KEYWORDS  
MGC.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
REFERENCE  
AUTHORS  
1 (bases 1 to 3635)  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
Scheetz,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S.,  
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
Abramson,R.D., Muljany,S.J., Bosak,S.A., McEwan,P.J.,  
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Fahey,J., Helton,E., Keltman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E.,  
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 3635)  
DIRECTOR MGC Project.  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCES  
DIRECT SUBMISSION  
Submitted (16-AUG-2004) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: http://mgc.nci.nih.gov  
Contact: MGC help desk  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Jeffrey Green M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: http://www.nisc.nih.gov/  
Contact: nisc\_mgc@hgrl.nih.gov  
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,

Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karins,E., Kwong,P., Latic,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantipop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAX Plate: 177 Row: a Column: 11  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23956207.

FEATURES

source

1..3635  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="MGC:90959 IMAGE:3597815"  
/tissue\_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."  
/clone\_lib="NCI CGAP\_Mam6"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..3635  
/gene="Ifih1"  
/note="synonyms: Helicard, Hlcd, MDA5, MDA-5"  
/db\_xref="GeneID:71586"  
/db\_xref="MGI:1918836"  
231..3161  
/gene="Ifih1"  
/codon\_start=1  
/product="Ifih1 protein"  
/protein\_id="AAH80200.1"  
/db\_xref="GI:51329784"  
/db\_xref="GeneID:71586"  
/db\_xref="MGI:1918836"

CDS

1  
/gene="Ifih1"  
/codon\_start=1  
/product="Ifih1 protein"  
/protein\_id="AAH80200.1"  
/db\_xref="GI:51329784"  
/db\_xref="GeneID:71586"  
/db\_xref="MGI:1918836"  
/translation="MSIVCSAEDSFRNLILFRRRLKMYIOVEPVLDELIFLSAETKE  
QILKINTCGNTSAEELISTLEQGWPLGWTOMFEVALEHSGNPLAARYVKTLDL  
PSPSETAHDECLHLTLPLQTLVDKLLINDVLDTCFEKGLTVEDNRISAAGNSG  
ESGVRELRLRIVQKENWSTFLDLRQTGNDALEQLTGGCPEDNTESDTSLAESV  
SCFDESIGHNSMGRDSTGMSDSBSVITQKRVSPBELQLRPYQMEVAQPALDKN  
IICLPTSGKTRVAAYITKDLDDKKQASEGKIVLVNKMALAEQLFRKFNPLYK  
KMYRIIGLSGDTQLKISPEVVKSYDVITISQILNLSLNLNLESGDDGVQLDSFLI  
IIECHHTNKEAVYNNIMRYLKOKLRNNDLKKONKPAIPLPOLGLTASPGVGAACK  
QSEAEKILINICANLDAFTIKTVKENLGQLKQIKPECKFVIADTRENPFKEKLE  
IMASIQTYCOKSPMSDFGTOYBOWAIOEKKAKDGNKRDVCAEHLRKYNALQIN  
DTRIMIDAYSHLETFTYDEKEKFAVLNDSDESDDEASSCNDQLKGVKSLKLDST  
EFLMNLFPDNKMKLKLAEKNPKYENKLIKLRNTILLEQTRSEBSRGIIFTKRQST  
YALSQWIMENAKFAEVGAHHLIGAGHSSEVPMTQEOKEVISKFRTEINLIAT  
TVABEGLDIKECNIVIRGLVTNEIAMVQARGARADESTYVLVTSSGSVTEREIVN  
DREKMYKAIINRVONMKPEEYAHKILEOVOSILEKMKVKRSIAKOYNDNPSLITL  
LCNKCMTLVCSEGNIVIEKMHVNMTPBFKGLIVRENKALQKRPADYQTNGEIICK  
CGQAMGTMMVHKGLDLPCLKIRNFVNPFKNSPKQYKKWVELPIRPDLDYSEYCLY  
SDSD"

ORIGIN

Alignment Scores:

Pred. No.: 6.8e-238 length: 3635  
Score: 4080.50 Matches: 800  
Percent Similarity: 84.4% Conservative: 66  
Best Local Similarity: 78.0% Mismatches: 109  
Query Match: 76.8% Indels: 51  
DB: 9 Gaps: 3

US-09-515-363C-2 (1-1025) x BC080200 (1-3635)

Qy 1 MetSerAenGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
Db 231 ATGTCGATTGCTGTTCTGCAGAGACAGCTTCAGGAATCATCTTATTCTTCAGGCC 290  
Qy 21 ArgVallyMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40

Db 291 AGGCTGAATAATGATCATTCAGGTGGAGCCAGTGTGACCACTCATCTTGTGTTCTGCA 350  
Qy 41 GluVallyGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 351 GAAACCAAGAAGACAGATTCTTAAAAAGATCAACACACTGTGTAAACACAGCGGCAGAA 410  
Qy 61 LeuLeuLeuSerThrLeuGluGlyValTTrpHisLeuGlyTTrpThrArgGluPheVal 80  
Db 411 CTGCTGTGAGCACCTTGAGAGCAGGACAGTGGCTCTGGGATGAGCGCAGATGTTCTGTG 470  
Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
Db 471 GAGCCCTAGACACAGTGGCAATCCCTTAGCCGCGCTATGTCAAAACCACTCATTT 530  
Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 531 GATCTGCCCTCTCCTCTGAGACTGCCCCATGACGAGTGTCTCCACTTGCTGACCCCTC 590  
Qy 121 LeuGlnProThrLeuValAspIlyLeuLeuValArgAspValLeuAspIlyCysMetGlu 140  
Db 591 CTCACGCCCACTTTGGTGGACAAACTTCTGATTACGATGTTGGACACTTGTCTCGAG 650  
Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
Db 651 AAGGACTCTTGACAGTGAAGACAGAAATCGATTCTGCTGACGAAACAGCGGGAAT 710  
Qy 161 GluSerGlyValArgGluLeuLeuGlyArgIleValGlnIlyGluAsnTrpPheSerAla 180  
Db 711 GAGTCAGGTGTAAAGAGAGCTGCTGAAGAAGATTGTCAGAAAGAAAATGGTTTCTACC 770  
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 771 TTCTGTGATGTTCTGCGCCAACTGGAATGATGCACTATTCAGAACTAACAGGTGGA 830  
Qy 201 AspCysSerGluSerAsnAlaGlnIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 831 GGCTGCCAGAAAGAAC----- 848  
Qy 221 GluGluGlnLeuLeuSerThrValGlnProAsnLeuGluIlyGluValTTrpGlyMet 240  
Db 848 ----- 848  
Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
Db 849 -----ACAGAAATCAGACACA 863  
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 864 AGTTGGCAGAAGAGTGTGAGCTGCTTCGATGAAGTCTTGACATTAACGCAACATG 923  
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnVal---AlaAlaArg 299  
Db 924 GGCAGGATTCAAGGACCATGGGAAGTGATTCAATGAAAGTGTCAATCCAGCAAAAAAGA 983  
Qy 300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319  
Db 984 GTATCCCCGAGCCAGAACTGCAAGTCAAGGCTTAACCAATGGAAGTGGCCCAACCACT 1043  
Qy 320 LeuGluGlyIlyAsnIleIleIleCysLeuProThrGlySerGlyIlySerThrArgValAla 339  
Db 1044 CTAGATGGGAAGAATATTATTATCTGCCCTCCCGGGAGTGGAAAAACAGAGTGGCT 1103  
Qy 340 ValTyrIleAlaIlyAspHisLeuAspIlyIlyIlyIlyAlaSerGluProGlyIlyVal 359  
Db 1104 GTTACATCAACCAAGATCACTTAGACAAGAGAGCAGCATCTGAATCCGGGAAGGTT 1163  
Qy 360 IleValIleuValAsnIlyValLeuLeuValGluGlnLeuPheArgIlyGluPheGlnPro 379  
Db 1164 ATCGTTCTGTCAATAAGTAATGTTAGCAGAACACTTTCCGAAAAGATTCAACCCA 1223  
Qy 380 PheLeuIlyIlyIlyTTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuIlyIleSer 399  
:::|||||

Db 1224 TATTGAAGAAATGSTATCGAATTATTGGATTAAAGTGGCGATACCAGCTGAATAATATCA 1283  
Qy 400 pheProGluValValIySerCysAspIleIleSerThrAlaGlnIleLeuGluAsn 419  
Db 1284 TTTCAGAAAGTTGTCAAAATCTTAACGATGTTATTATGACGACTGCTCAAAATCCTTGAAAAAC 1343  
Qy 420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439  
Db 1344 TCCCTCTTAAATCTGAGAGTCGAGACGATGACGGTGTGACGCTGTACAGACTTCTCTCTC 1403  
Qy 440 IleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArg 459  
Db 1404 ATTATCATTTGATGAGTGCATCACACCAAGAGGCGAGTCTATAACAACATCATGAGA 1463  
Qy 460 HisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysProValIle 479  
Db 1464 CGATATTGGAAGCAGAGCTGAGAAACAATGACCTCAAGAAACAACCAAGCCATT 1523  
Qy 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGln 499  
Db 1524 CCCCTGCCGAGATAGTACTAGACTGACAGCTTCACTGCTGTTGAGCAGCCAAAAAGCAG 1583  
Qy 500 AlaLysAlaGluGlnHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519  
Db 1584 TCTGAGCGTGAATAAACATATTTAAATATATGTGCCAATCTTGATGCCCTTACATTAA 1643  
Qy 520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnIleProCysLysLysPhe 539  
Db 1644 ACAGTGAAGAAGAAATCTTGTCAACTCAAAACACCAATTAAGAACCATGCAAGAAATTT 1703  
Qy 540 AlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGlnIleMetThr 559  
Db 1704 GTGATTGCTGATGACACACAGAAATCCATTAAAGAGAAACTTCTAGAATTAATGGCA 1763  
Qy 560 ArgIleGlnThrTyrCysGlnMetSerPrometSerAspPheGlyThrGlnProTyrGlu 579  
Db 1764 AGCATTCAGACTTACTGCCAAAAAAGTCCAAATGTCAGATTTTGGAAACCCCAACTTATGAG 1823  
Qy 580 GlnTrrAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgVal 599  
Db 1824 CAGTGGCCCATTCAAATGAGAAAAAAGCTGCTAAAGACGGAATTCGAAAGATCGCGTC 1883  
Qy 600 CysAlaGlnHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619  
Db 1884 TGTGAGAAACATTTGAGGAAGTCAACGAAAGCCCTACAAATCAACGACACGATCCGAATG 1943  
Qy 620 IleAspAlaTyrThrHisLeuGlnThrPheTyrAsnGlnLysAspLysPheAla 639  
Db 1944 ATTGATGCATATAGCCACCTGGAGACATTTCTACACTGATGAGAAAGAAAGATTCCGCA 2003  
Qy 640 ValIleGluAspAspSerAspGluGlyLysAspAspGluTyrCysAspGlyAspGluAsp 659  
Db 2004 GTCCCTC--AATGACAGCGACGAGATGATGACGAGGCCACAGTTGCAATGACCAACTT 2060  
Qy 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679  
Db 2061 AAGGCGCATGTAAAGAAATCTTGAACCTGACGAAACGATGAATTTCTCATGAATTG 2120  
Qy 680 PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu 699  
Db 2121 TTCTTTGATTAACAAAGAAATGTTGAAAAAACTAGCTGAAAAACCAAAATACGAAATGAA 2180  
Qy 700 LysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluSerAla 719  
Db 2181 AAACCTCAATAATTAAAGAAACAGATACTGAAACAATTCACAAGGTCTGAGAGTCCCTCC 2240  
Qy 720 ArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrrIleThr 739  
Db 2241 CGAGGAATTAATTTTACAAAAACAGACAGAGCACCTACGCACTTCCCACTGATCATG 2300  
Qy 740 GluAsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHis 759  
Db 2301 GAAATGCAAAAGTTTGGGGAAGTTGAGAGTCAAAAGCGCATCACCTGATTGGCGCGGCGAC 2360

Qy 760 SerSerGluPheLysPrometThrGlnAsnGluGlnLysGluValIleSerLysPheArg 779  
Db 2361 AGCAGTGAAGTCAAGCCCACTGACTCAGACTGAAACAAAAAGAAAGTCATTAGTAATTTCCG 2420  
Qy 780 ThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys 799  
Db 2421 ACTGGCGAAATAAATCTGCTTATCGCTACGACGGTGGCAGAGGAAGCCTGGATATCAAA 2480  
Qy 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819  
Db 2481 GAGTCAATATATTGTTATTCGTTATGGCTTGTCAGAACGAGATAGCCATGGTCCAGGCC 2540  
Qy 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839  
Db 2541 CGGGGTCCAGCCAGAGCTGATGAAAGCACGTAATGCTTGTCACCAAGCAGTGGCTCAGGA 2600  
Qy 840 ValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859  
Db 2601 GTTACCGAACGGAGATTGTTAATGATTTCCGAGAAAGATGATGATATAAGCTATTAAAC 2660  
Qy 860 CysValGluAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGln 879  
Db 2661 CGTGTCAAAACATGAACCAAGAGATGACACATAAGATTTTGGAAATGCAAGTGCACAA 2720  
Qy 880 SerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsn 899  
Db 2721 AGTATCCTGGAAGAAAGAAATGAAGTCAAAAGAACATTCGCAAGCAATGCAACGACAT 2780  
Qy 900 ProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAsp 919  
Db 2781 CCATCGTTAATTAACACTTCTCTGCCAAAAATTTGTAGCATGCTGCTGCTCGGAGAAAC 2840  
Qy 920 IleHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyr 939  
Db 2841 ATCCATGTCAATTGAGAAAGATGCATCATGTCAATATGACACCAAGAAATTCAGAGCACTTAC 2900  
Qy 940 IleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlu 959  
Db 2901 ATTGTAGAGAAACAAAGCACTGCAAAAGAAATTTGCTGATTATCAGACCACAAATGGAAG 2960  
Qy 960 IleIleCysLysCysGlyGlnAlaTrrGlyThrMetMetValHisLysGlyLeuAspLeu 979  
Db 2961 ATTATCTGCAAGTGTGGCCAGGCTTGGGAACAATGATGTGCACAAAGTTTAGATTGG 3020  
Qy 980 ProCysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGln 999  
Db 3021 CCTTGCTTAAATAAGCAATTTTGTAGTCAATTTCAAAAATACTCACCGAAGAAACAG 3080  
Qy 1000 TyrLysLysTrrValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019  
Db 3081 TACAAGAAAGTGGTGAATTCCTATCAGATTTCTGATCTTGACTACTCAGAAATACTGC 3140  
Qy 1020 LeuPheSerAspGluAsp 1025  
Db 3141 TTGTATAGTATGAAGAT 3158  
RESULT 15  
BC078180 2540 bp mRNA linear PRI 27-JUL-2004  
LOCUS BC078180  
DEFINITION Homo sapiens interferon induced with helicase C domain 1, mRNA  
(CDNA clone IMAGE:6502757), partial cds.  
ACCESSION BC078180  
VERSION BC078180.1 GI:50415794  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.  
REFERENCE 1 (bases 1 to 2540)  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haleh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Boufard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932  
2 (bases 1 to 2540)  
Strausberg,R.  
Direct Submission  
Submitted (19-JUL-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK  
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)  
Akhter,N., Ayele,K., Beckettrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Boufard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Lalic,P., Legaapi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantiripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAK Plate: 174 Row: p Column: 16  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27886567.

FEATURES  
source  
1. .2540  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:502757"  
/tissue\_type="uterus, leiomyosarcoma"  
/clone\_lib="NIH MGC\_71"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1. .>2540  
/gene="IFIH1"  
/note="synonyms: MDA5, MDA-5, Hicd"  
/db\_xref="GeneID:64135"  
/db\_xref="MIM:606951"  
224. .>2540  
/gene="IFIH1"  
/codon\_start=1  
/product="IFIH1 protein"  
/protein\_id="AAH78180.1"  
/db\_xref="GI:50415795"

CDS  
gene

/db\_xref="GeneID:64135"  
/db\_xref="MIM:606951"  
/translation="MSNGYSTDENFRYLISCFRARVKMYIOVEPLDYLTPLPAEVKE  
QIQRTVATSGNQAVEILLSTLEKGVHIGWTRFEALRRTPGSLARYMNPETDL  
PSPFENAHDEYQLNLLOPTLVKLVKRDVLDKMEBELTIEDNRNIAAENNGN  
ESGVRELKRIYOKENWFSAFNLVLRQTNNELVQELTSGDCSESNAIENLSQVDP  
QVEQLSTVQPNLEKEVMGNNSSSPADSVSESDTSLAEGSVSLDESIG  
NSNMGSDSGTMSGDSDEENVAARAPPELQRPYOMEVAQPLAEGKNIICLPITSG  
KTRVAVYIAKHLDKKKASEPGKIVLVNKVLVLEQLFRKEPQPLKMYRVLIGSG  
DTOLKISFPEVVSCTIISTAQILNSLNLNENGEDAGVQLSDFSLIIDECHTNK  
EAVYNNIMRHYLMQKLKNRLKKNKPIVLPQILGLTASPGVGATKQAKAEHLK  
LCANLDAFTIKTYKENLDOLKNQIOEPCKPAIADATREDPFKEKLEIMTRIQTCC  
MSPMSDFGTOPYEOMAIOMEKKAKEGNRKRVCAEHLRKYNELQINDTIRMIDAYT  
HLETFYNEEKOKKFAVIDEDSDGDEYCDGEDDEDLKKPLKDETDRFLMTLFFE  
NNKMLKRLAENPEYENKLTCLRNTIMEQYTRTEESARGIIFTRQSAAYALSOWITE  
NEKFAEVGVKAHHLIGAGHSSEFKPMQKKK"

ORIGIN

Alignment Scores:

Pred. No.:	2.15e-230	Length:	2540
Score:	3954.00	Matches:	768
Percent Similarity:	99.9%	Conservative:	3
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	74.4%	Indels:	0
DB:	8	Gaps:	0

US-09-515-363C-2 (1-1025) x BC078180 (1-2540)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
DB 224 ATGTGAAATGGGTAATCCACAGACGAGAAATTCCGCTATCTCATCTCGTCTCAAGGCC 283

QY 21 ArgValIysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
DB 284 AGGGTGAATAATGATCATCAGGTGAGCCTGTGCTGGACTAAGCTTCTGCTGCA 343

QY 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
DB 344 GAGGTGAAGAGACAGATTGACAGACAGTCGCCACTCCGGAAACATGACGCAATTGAA 403

QY 61 LeuLeuLeuSerThrLeuGluIysGlyValTropHisLeuGlyTyrThrArgGluPheVal 80  
DB 404 CTGCTGCTGAGCACCTTGAGAAAGGAGTCTGGCACCCTTGTTGACTCGGGAATTCGTG 463

QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
DB 464 GAGGCCCTCCGAGAACCCGACGCCCTCTGGCCGCCCTACATGAACCTTGAGCTCACG 523

QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
DB 524 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCAATGCTGAACTC 583

QY 121 LeuGlnProThrLeuValAspIysLeuLeuValArgAspValLeuAspIysCysMetGlu 140  
DB 584 CTTGAGCCCACTCTGGTGCAAGCTTCTAGTTAGAGACGTCCTTGATTAAGTCATGAG 643

QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaIaGluAsnAsnGlyAsn 160  
DB 644 GAGGAAGTGTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAAATGGAAT 703

QY 161 GluSerGlyValArgGluLeuLeuIysArgIleValGlnIysGluAsnTrpPheSerAla 180  
DB 704 GAATCAGGTGTAGAGAGCTAATAAAAGGATTGTGCAGAAAGAAAAGTGTCTCTGCA 763

QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
DB 764 TTTTCTGAATGTTCTTCTGTAACAAGAAACAAATGAACCTTGTCCAAGAGTTAAACAGGCTCT 823

QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
DB 824 GATTGCTCAGAAAGCAATGACAGAGATTGAGAAATTATCACAAAGTTGATGCTCTCAAGTG 883

QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIysGluValTyrGlyMet 240

Db 884 GAAGCAACTCTTCAACCAAGTTCAGCCAAATCTGAGAAGAGGTCTGGGCATG 943  
Qy 241 GluAenAsnSerSerGluSerSerPheAlaAspSerSerValSerGluSerAspThr 260  
Db 944 GAGAAATACTCATCAGAAATCTTTGGAGATTCTTCTGTAGTTTCAGAAATCAGACACA 1003  
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 1004 AGTTTGCAGAGAAGAGTGTCACTGCTTATGATGAAGCTTTGGACATAACAGCAACATG 1063  
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
Db 1064 GGCAGTGATTCAAGCACCATGGAGAGTGATTCAGATGAAAGAAATGTGGCAGCAAGACA 1123  
Qy 301 SerProGluProGluLeuGluLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
Db 1124 TCCCCGAGCCAGAACTCCAGCTCAGGCCCTTACCAAATGGAAGTTGCCCAGCCAGCCTTG 1183  
Qy 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340  
Db 1184 GAAGGAGAATATCATCATCTGCTCCCTACAGGAGTGGAAGAAACAGAGTGGCTGTT 1243  
Qy 341 TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle 360  
Db 1244 TACATTGCCCAAGGATCATCTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAGTTATA 1303  
Qy 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380  
Db 1304 GTTCTTGTCAATAAGGTACTGTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACTATT 1363  
Qy 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400  
Db 1364 TTGAAGAAATGTATCGTGTATTGATTAAAGTGTGATACCCAACTGAAATATCATTT 1423  
Qy 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
Db 1424 CCAGAGTTGTCAAGTCCGTGTATTTATTTATCAGTACAGCTCAAAATCCTGAAAACTCC 1483  
Qy 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
Db 1484 CTCTTAACTTGGAATGGAATGAGAGATGCTGTGTTCAATTGTCAGACTTTCCCTCATTT 1543  
Qy 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460  
Db 1544 ATCATTTGATGAATGTCTATCACACCAACAAGAGCAGTATTAATAACATCATGAGGCAT 1603  
Qy 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480  
Db 1604 TATTTGATGCAGAGTTGAAACAAATAGACTCAAGAAAGAAACAAACCAAGTATTCCTCC 1663  
Qy 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500  
Db 1664 CTTCTCAGATACTGGGACTTAACACTTCACTGTGTGGAGGGGCCAGAGCAAGCC 1723  
Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
Db 1724 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCAATTTACTATTAAAACT 1783  
Qy 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540  
Db 1784 GTTAAAGAAAACCTTGATCAACTGAAAAAACCAATACAGGAGCCATGCAAGAGTTTGC 1843  
Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560  
Db 1844 ATTGCAGATGCAACAGAGAAAGATCAATTAAGAGAACTTTCTAGAAATATGACAAAG 1903  
Qy 561 IleGlnThrTyrCysGlnMetSerPrometSerAspPheGlyThrGlnProTyrGluGln 580  
Db 1904 ATTCAAACTTATGTCAAAATGAGTCAATGTCAGATTTTGGAACTCAACCTATGAACAA 1963  
Qy 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600

Db 1964 TGGGCCATTCAATGGAAGAAAAAGCTGCAGAAAGAGAAATCGCAAGAACGTGTTGT 2023  
Qy 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
Db 2024 GCAGAACATTTGAGGAAGTACAAATGAGGCCCTTACAAATTAATGACACAACTTGAATGATA 2083  
Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640  
Db 2084 GATGCGTATACTCATCTTGAAACTTTCTATTAATGAAAGAAAGTAAAGATTGGCAGTC 2143  
Qy 641 IleGluAspAspSerAspGluGlyLysAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 2144 ATAGAAATGATAGTATGATGAGGGTGTGATGATGATGATGATGATGATGATGATGAG 2203  
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2204 GATGATTTAAAGAAACCTTGAAACTGGATGAACAGATAGATTCTCATGACTTTATT 2263  
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2264 TTTGAAAAACAATAAATGTGGAAGAGCTGGCTGGAAGAACCAAGATATGAAGAAAG 2323  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720  
Db 2324 CTGACCAATTTAAGAAATACCATTAATGAGCAATATATACTAGACTGAGGAATCAGACGA 2383  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740  
Db 2384 GGAATTAATCTTTACAAAAACAGACAGAGTGCATATGCCCTTCCAGTGAATTACTGAA 2443  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760  
Db 2444 AATGAAAAATTTGCTGAAGTGAAGTCAAGGCCCACTGATTTGAGCTGGAACACAGC 2503  
Qy 761 SerGluPheLysPrometThrGlnAsnGluGlnLys 772  
Db 2504 AGTGAGTTCAAAACCATGACACAAAAA 2539

Search completed: March 13, 2006, 18:53:20  
Job time : 12742 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 12, 2006, 22:09:59 ; Search time 2412 Seconds  
(without alignments)  
2832.218 Million cell updates/sec

Title: US-09-515-363C-2

Perfect score: 5311  
Sequence: 1 MSGNGYSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSDDED 1025

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp  
-Q=/absas/ABSSWEB.spool/US09515363/runat\_10032006\_181127\_7925/app\_query.fasta\_1  
-DB=N Geneseg -QFWT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blissum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=absas02p  
-USER=US09515363@CCN\_1\_1\_1929@runat\_10032006\_181127\_7925 -NCPU=6 -ICPU=3  
-NO\_MMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseg\_21:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002as:\*
- 7: geneseqn2002bs:\*
- 8: geneseqn2003as:\*
- 9: geneseqn2003bs:\*
- 10: geneseqn2003cs:\*
- 11: geneseqn2003ds:\*
- 12: geneseqn2004as:\*
- 13: geneseqn2004bs:\*
- 14: geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match	Length	DB	ID	Description
No.						
1	5311	100.0	3365	4	AAD17203	Aad17203 Human mel
2	5311	100.0	3380	12	ADJ74904	Adj74904 Marker ge
3	5311	100.0	3380	13	ADR14404	Adr14404 Human NF-
4	5311	100.0	3380	14	AEA23691	Aea23691 Human PRO

5	5311	100.0	3380	14	AEA36113	Aea36113 Human nuc
6	5307	99.9	3434	14	ADY17563	Ady17563 DNA encod
7	5307	99.9	3434	14	ADY20376	Ady20376 DNA encod
8	5303	99.8	3928	11	ACN91892	Acn91892 Breast ca
9	5285	99.5	3372	6	ABA04908	Abao04908 Human RNA
10	5285	99.5	3446	10	ADC30823	Adc30823 Human nov
11	5283	99.5	3379	12	ADN04879	Adn04879 Antipsori
12	5276.5	99.4	3131	4	AAD11170	Aad11170 Human mel
13	5251	98.9	3668	12	ADQ22237	Adq22237 Human sof
14	4194	79.0	3771	12	ADJ75813	Adj75813 Marker ge
15	3176	59.8	1967	4	AAS40960	Aas40960 CDNA enco
16	2874	54.1	1776	6	ADA53260	Ada53260 Human cod
17	2442	46.0	1443	6	ABA04916	Abao04916 Human RNA
18	2356.5	44.4	1557	4	AAI59285	Aai59285 Human pol
19	2120	39.9	1382	4	AAI61071	Aai61071 Human pol
20	2120	39.9	1382	10	ADC32572	Adc32572 Human nov
21	2072	39.0	1392	11	ADI30722	Adi30722 Human CDN
22	2072	39.0	1392	13	ADS82789	Ads82789 Human lym
23	1995	37.6	1284	6	ABA04913	Abao04913 Human RNA
24	1961.5	36.9	1319	4	AAS41517	Aas41517 CDNA enco
25	1352	25.5	2590	6	ABZ11889	Abz11889 Human pol
26	1352	25.5	2590	12	ADM44407	Adm44407 Novel hum
27	1344	25.3	2613	4	AAH15201	Aah15201 Human CDN
28	1201.5	22.6	1258	4	AAS01149	Aas01149 Interfero
29	1201.5	22.6	1270	4	AAS01151	Aas01151 Interfero
30	1129	21.3	2250	11	ADM03387	Adm03387 Human CDN
31	1106.5	20.8	956	5	AAS91688	Aas91688 DNA encod
32	1091.5	20.6	4372	13	ADR97293	Adr97293 Human RIG
33	1091.5	20.6	4372	14	ADY17515	Ady17515 DNA encod
34	1091.5	20.6	4372	14	AEB22799	Aeb22799 Human col
35	1090.5	20.5	3065	10	ADF76334	Adf76334 Novel hum
36	1090.5	20.5	3065	12	ADJ74908	Adj74908 Marker ge
37	1090.5	20.5	3065	12	ADJ75029	Adj75029 Marker ge
38	1090.5	20.5	3065	12	ADN04859	Adn04859 Antipsori
39	1090.5	20.5	3065	13	ADR25611	Adr25611 Breast ca
40	1090.5	20.5	3065	13	ADR14232	Adr14232 Human NF-
41	1090.5	20.5	3065	13	ADP25010	Adp25010 PRO polyp
42	1090.5	20.5	3065	14	AEA23524	Aea23524 Human PRO
43	1090.5	20.5	3065	14	AEA36112	Aea36112 Human nuc
44	1090.5	20.5	3258	8	ABX63631	Abx63631 Human CDN
45	1069	20.1	4468	13	ACN41904	Acn41904 Human dia

ALIGNMENTS

RESULT 1	
AAD17203	
ID	AAD17203 standard; CDNA; 3365 BP.
XX	
AC	AAD17203;
XX	
DT	29-NOV-2001 (first entry)
XX	
DE	Human melanoma differentiation associated (Mda)-5 CDNA.
XX	
KW	Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;
KW	RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
KW	neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
KW	breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
KW	central nervous system; cytosolic; apoptosis; ss.
XX	
OS	Homo sapiens.
XX	
FH	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	169..3246
FT	/*tag= a
FT	/product= "Human melanoma differentiation associated
FT	(Mda)-5 protein"
FT	3225..3228
FT	/*tag= b
FT	/note= "ATTTA motif"
FT	3284..3287
FT	/*tag= c
FT	misc_feature

FT /note= "ATTTA motif"  
FT polyA\_signal 3343..3348  
FT /\*tag= d  
XX WO200164707-A1.  
XX PN  
XX 07-SEP-2001.  
PD  
XX 28-FEB-2001; 2001WO-US006960.  
XX  
XX 29-FEB-2000; 2000US-00515363.  
XX  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX  
XX PI Fisher PB, Kang D, Gopalkrishnan RV;  
XX  
XX WPI; 2001-565494/63.  
XX P-PSDB; AAE10155.  
XX  
XX Nucleic acid sequences encoding a Melanoma Differentiation Associated  
PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral  
PT activity.  
XX  
XX PS Claim 1; Page 16-18; 152pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid encoding a  
CC melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5  
CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.  
CC Mda-5 is a novel interferon (IFN) inducible gene with structural  
CC similarities to RNA helicases and CARD motif containing proteins. Mda-5  
CC is induced during terminal differentiation in human melanoma cells  
CC treated with the combination of recombinant fibroblast IFN and the  
CC antileukaemic compound mezerein (MEZ). Mda-5 is useful for identifying  
CC compounds that may induce its expression. Mda-5 is useful for treating  
CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma  
CC multiforme, cervical cancer, breast cancer, colon cancer, prostate  
CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a  
CC cancer of the central nervous system and apoptosis. The Mda-5 promoter  
CC exhibits melanocyte tissue specificity and minimises systemic toxicity.  
CC The present sequence is human Mda-5 cDNA  
XX  
XX SQ Sequence 3365 BP; 1138 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0 Length: 3365  
Score: 5311.00 Matches: 1025  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 4 Gaps: 0

US-09-515-363C-2 (1-1025) x AAD17203 (1-3365)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
DB 169 ATGTCGAATGGGTATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGCTTCAGGGCC 228  
QY 21 ArgValIlyMetTyrIleGlnValGluProValIleuAspTyrLeuThrPheLeuProAla 40  
DB 229 AGGGTGAATAATGATACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTGCTGCA 288  
QY 41 GluValIlyGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
DB 289 GAGGTGAAGACGACGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAAGGCAATTGAA 348  
QY 61 LeuLeuLeuSerThrLeuGluIlyGlyValTPrHisLeuGlyTyrThrArgGluPheVal 80  
DB 349 CTGCTGTGAGCACCTTGAGAGGAGTCTGGCACCTTGCTGACTCGGGAATTGCTG 408  
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
DB 409 GAGGCCTCCGGAGAACCGGACGCTCTGGCCGCCCGCTAACATGAACCTGAGCTCACG 468

QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
DB 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCAAGCTGCTGAACCTC 528  
QY 121 LeuGlnProThrLeuValAspIlySLeuLeuValArgAspValIleuAspIlyS CysMetGlu 140  
DB 529 CTTCAAGCCCACTCTGGTGACAAGCTTCTAGTTAGAGACGCTCTTGATTAAGTCATGGAG 588  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
DB 589 GAGGAAGTGTGAACAATGAAGACAGAAACCGGATTCGCTGCAGAAAACAATGGAAT 648  
QY 161 GluSerGlyValArgGluLeuLeuIlySArgIleValGlnIlyS GlnAsnTrpPheSerAla 180  
DB 649 GAATCAGGTGTAAAGAGAGCTACTAAAAAGGATTGCAAGAAAGAAAAGTGTCTCTGCA 708  
QY 181 PheLeuAsnValIleuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
DB 709 TTTCTGAATGTTCTTCGTCAACAGAAACAATGAATCTGCCAAGAGTTAACAGGCTCT 768  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
DB 769 GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTG 828  
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIlyS GlnValTPrGlyMet 240  
DB 829 GAAGGCAACTCTTTTCAACACAGATTCAAGCAAATCTGGAGAAAGAGTCTGGGGCATG 888  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValIleSerGluSerAspThr 260  
DB 889 GAGAAATACTCATCAGAAATCATCTTTTGAGATTCTCTGTAGTTTCAGAAATCAGACACA 948  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
DB 949 AGTTGGCAGAGAGAGAGTGCAGCTGCTTAGATGAAGAAGTCTTGACATTAACAGCAACATG 1008  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
DB 1009 GGCAGTGATTCAAGCACCATGGGAAGTGAATTCAGATGAAGAGATGTGGCAGCAAGACA 1068  
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
DB 1069 TCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCCGCAGCCTTG 1128  
QY 321 GluGlyIlyAsnIleIleIleCysLeuProThrGlySerGlyIlySerThrArgValAlaVal 340  
DB 1129 GAAGGGAAGATATCATCATCTGCCCTCCACAGGAGTGAAAAACAGAGTGGCTGTT 1188  
QY 341 TyrIleAlaIlyAspHisLeuAspIlySLeuIlySLeuAlaSerGluProGlyIlySValIle 360  
DB 1189 TACATGCCAAGGATCATTAGACAAGAAAGAAACATCTGAGCCTGGAAGATTATA 1248  
QY 361 ValLeuValAsnIlySValLeuLeuValGluGlnLeuPheArgIlyS GlnPheGlnProPhe 380  
DB 1249 GTTCTGTCAATAAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAACCATTT 1308  
QY 381 LeuIlySLeuTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuIlySLeuPhe 400  
DB 1309 TTGAAGAAATGCTATCGTGTTATTGATTAAAGTGATACCCCACTGAATAATATCATTT 1368  
QY 401 ProGluValValIlySLeuSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
DB 1369 CCAGAGTTGTCAAGTCCGTGATATTATTATATCAGTACAGCTCAAAATCCTTGAAGATCC 1428  
QY 421 LeuLeuAsnLeuGluAsnGlyIlySAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
DB 1429 CTCTTAACCTTGAAAAATGGAGAGATGTGTTCAATTGTCAAGCTTTTCCCTCATTT 1488  
QY 441 IleIleAspGluCysHisHisThrAsnIlyS GlnAlaValTyrAsnAsnIleMetCysArgHis 460  
DB 1489 ATCATTTGATGATGTATCATCACCAACAAGAGCAGTGTAATAATAACATCATGAGGCAT 1548  
QY 461 TyrLeuMetGlnIlySLeuIlySAsnAsnArgLeuLeuIlyS GlnAsnIlySProValIlePro 480

Db 1549 TATTGATGCAGAGTTGAAAAACATAGACTCAAGAAAGAAAAACCAAGTATTCCC 1608  
Qy 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500  
Db 1609 CTTTCCTCAGATACTGGGACTTAACAGCTTCACCTGGTGTGAGGGGCCACGAAGCAAGCC 1668  
Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
Db 1669 AAAGCTGAAGAACACATTTTAAACTATAGTGCCTCAATCTTGATGCATTACTATTAAAACT 1728  
Qy 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540  
Db 1729 GTTAAAGAAAACTTGATCAACTGAAAAACCAATATACAGACCATGCAAGAGTTTGCC 1788  
Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560  
Db 1789 ATTGCAGATGCCAACGAGAGATCCATTAAAGAGAACTTCTAGAATAATGACAAAG 1848  
Qy 561 IleGlnThrTyrcysGlnMetSerPrometSerAspPheGlyThrGlnProTyrcGluGln 580  
Db 1849 ATTCAAACTTATTGTCAATGAGTCCAATGTCAGATTTTGGAACCTCAACCTATGAACAA 1908  
Qy 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
Db 1909 TGGGCATTCAAAATGAAAAAAAGCTGCAAAAAAAGAAATCGCAAGAACGTGTGT 1968  
Qy 601 AlaGluHisLeuArgLysTyrcAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
Db 1969 GCAGAACATTTGAGGAAGTACATGAGGCCCTCACAAAATTAATGACACAAATTCGAATGATA 2028  
Qy 621 AspAlaTyrrThrHisLeuGluThrPheTyrrAsnGluGluLysAspLysLysPheAlaVal 640  
Db 2029 GATGCCATACTCATCTTGAACCTTCTATATATGAGAGAGAAAGATGCAAGTTCGACATC 2088  
Qy 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrcysAspGlyAspGluAspGlu 660  
Db 2089 ATAGAACATGATAGTGATGAGGGTGTGATGATGAGTATTGTGATGTGATGATGAGATGAG 2148  
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2149 GATGATTTAAAGAAACCTTTGAAACTGATGAAACAGATGATTCTCATGACTTTATTT 2208  
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrcGluAsnGluLys 700  
Db 2209 TTTGAAAAACATAAATGTTGAAAAAGCTGGCTGAAAAACCAGAATATGAAAAAG 2268  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrrThrArgThrGluLysSerAlaArg 720  
Db 2269 CTGACCAATTAAGAAATACCAATATGAGCAATATACTAGACTGAGGAATCAGCAGCA 2328  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrrAlaLeuSerGlnTrpIleThrGlu 740  
Db 2329 GGAATTAATCTTTACAAAAACAGACAGAGTGCATATGCGCTTCCAGTGGATTACTGAA 2388  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisIleLeuIleGlyAlaGlyHisSer 760  
Db 2389 AATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGAGCTGACACAGC 2448  
Qy 761 SerGluPheLysPrometThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2449 AGTGAAGTTCAAAACCATGACACAGATGAACAAAAAGAGTCATTAGTAAATTTCGACT 2508  
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2509 GGAATAATCAATCTGCTTATCGCTACACAGTGCAGAGAGAGTCTGATATTTAAAGAA 2568  
Qy 801 CysAsnIleValIleArgTyrcGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2569 TGTAAACATTTTATCCGTATGTGTCTGTCCACCAATGAATAAGCATGTGTCCAGGCCGT 2628  
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrrValLeuValAlaHisSerGlySerGlyVal 840

Db 2629 GGTGAGCCAGAGCTGATGAGACACCTACGTCCTGCTTGCTCACAGTGTTCAGAGTT 2688  
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrrLysAlaIleHisCys 860  
Db 2689 ATGAAACATGAGACAGATTGAATGATTTCCGAGAGAAAGATGATGTATTAAGACTATACATTGT 2748  
Qy 861 ValGlnAsnMetLysProGluGluTyrrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2749 GTTCAAAATATGAACCCAGAGAGTATGCTCATTAAGATTTTGAATTTACAGATGCAAGT 2808  
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrrLysAsnAsnPro 900  
Db 2809 ATATGGAAGAAAGAAATGAAGAACCAAGAGAAATATGCAAGCATTAAGAAATTAACCCA 2868  
Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920  
Db 2869 TCACTAATACTTCTTCCTTGCAAAAACTGCAGTGTGCTTAGCCTGTTCTGGGGAAGATATC 2928  
Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrrIle 940  
Db 2929 CATGTAATTGAGAAATGATCATCAGTCAATATGACCCAGAAATTCAGAGAACTTTACATT 2988  
Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrcGlnIleAsnGlyGluIle 960  
Db 2989 GTAAGAGAAAAACAAAGCATGCAAAAAAGAGTGTGCCACTATCAATTAATGTGAAATC 3048  
Qy 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3049 ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTCACAAAGGCTTAGATTGGCCT 3108  
Qy 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrr 1000  
Db 3109 TGTCTCAAAATTAAGGAATTTGTAGTGTGTTTCAAAAATTAATTCACAAAGAAACATATC 3168  
Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrrSerGluCysCysLeu 1020  
Db 3169 AAAAAGTGGTAGAATTACTATCACATTTCCAAATCTTGACTAATTGAGAAATGCTGTTTA 3228  
Qy 1021 PheSerAspGluAsp 1025  
Db 3229 TTTAGTGATGAGGAT 3243  
RESULT 2  
ADJ74904  
ID ADJ74904 standard; DNA; 3380 BP.  
XX  
AC ADJ74904;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Marker gene SEQ ID NO:156.  
XX  
KW bronchial asthma; chronic obstructive pulmonary disease;  
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;  
KW gene therapy; marker gene; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN EP1394274-A2.  
XX  
PD 03-MAR-2004.  
XX  
PF 04-AUG-2003; 2003EP-00254857.  
XX  
PR 06-AUG-2002; 2002JP-00229312.  
PR 20-MAR-2003; 2003JP-00077212.  
XX  
PA (GENO-) GENOX RES INC.  
XX  
PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;  
XX  
DR WPI; 2004-193155/19.  
XX

PT Testing for bronchial asthma or chronic obstructive pulmonary disease by  
PT comparing the expression level of a marker gene in a biological sample  
PT from a subject with the expression level of the gene in a sample from a  
PT healthy subject.  
XX  
PS Claim 1; SEQ ID NO 156; 241pp; English.  
XX

CC The present invention describes a method of testing for bronchial asthma  
CC or chronic obstructive pulmonary disease. The method comprises  
CC determining the expression level of a marker gene in a biological sample  
CC from a subject, comparing the expression level determined with the  
CC expression level of the marker gene in a biological sample from a healthy  
CC subject, and judging whether the subject has bronchial asthma or chronic  
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of  
CC genes (S1) whose expression levels increase when respiratory epithelial  
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)  
CC whose expression levels decrease when respiratory epithelial cells are  
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for  
CC testing for bronchial asthma or chronic obstructive pulmonary disease;  
CC (2) a kit for screening for a candidate compound for a therapeutic agent  
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)  
CC an animal model for bronchial asthma or chronic obstructive pulmonary  
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a  
CC method for producing an animal model for bronchial asthma or chronic  
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial  
CC asthma or chronic obstructive pulmonary disease, comprising the compound,  
CC a marker gene or an antisense nucleic acid corresponding to a portion of  
CC the marker gene, a ribozyme, a polynucleotide that suppresses the  
CC expression of the gene through an RNAi effect or an antibody recognising  
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for  
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a  
CC probe has been immobilised to assay a marker gene. (I) has respiratory  
CC and antiasthmatic activities, and can be used in gene therapy. The method  
CC is useful for testing for or screening for a therapeutic agent for  
CC bronchial asthma or chronic obstructive pulmonary disease. The present  
CC sequence is used in the exemplification of the present invention.

XX Sequence 3380 BP; 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0 Length: 3380  
Score: 5311.00 Matches: 1025  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 12 Gaps: 0

US-09-515-363C-2 (1-1025) x ADJ74904 (1-3380)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
DB 169 ATGTCGAATGGGTATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGCCTCAGGGCC 228  
QY 21 ArgValIysMetTyrIleGlnValGluProValIleuAspTyrLeuThrPheLeuProAla 40  
DB 229 AGGGTGAATAATGTACATCCAGGTGGAGCCTGTGCTGACTGACCTGACCTTTCGCTGCA 288  
QY 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
DB 289 GAGGTGAAGAGACAGATTCAAGAGACAGTCGCCACCTCCGGGAACATGACGCGACTTGA 348  
QY 61 LeuLeuLeuSerThrLeuGluIlyGlyValTyrPheIleuGlyTyrThrArgGluPheVal 80  
DB 349 CTGCTGTGAGCACCTTGAGAAAGGAGTCTGGCACCTTGCTTGACTCGGGAATTTCGTG 408  
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
DB 409 GAGGCCCTCCGAGAACCGGACGCTCTGGCCGCCCTTACATGAACCTGAGCTCAGC 468  
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
DB 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 528

QY 121 LeuGlnProThrLeuValAspIlyLeuLeuValArgAspValIleuAspIlySerMetGlu 140  
DB 529 CTTCAGCCCACTCTGTGGACAAAGCTTCTAGTGAAGACGCTTGATGAATGATGATGAG 588  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaIleGluAsnAsnGlyAsn 160  
DB 589 GAGGAACCTGTGACAAATTGAAGACAGAAACCAGATTGCTGCTGCAGAAAAACAATGGAAT 648  
QY 161 GluSerGlyValArgGluLeuLeuLeuIlyAspArgIleValGlnIlyGluAsnTrpPheSerAla 180  
DB 649 GAATCAGGTGTAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGTTCTCTGCA 708  
QY 181 PheLeuAsnValIleuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
DB 709 TTTCTGAATGTTCTTCGTCAAACAGAAACAAATGAACCTGTCCAAGAGTTAAACAGGCTCT 768  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
DB 769 GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTATCACAAGTTGATGGTCCCAAGTG 828  
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIlyGlyValTyrGlyMet 240  
DB 829 GAAGAGCAACTCTTTTCAACCAACAGTTCAGCCAAATCTGGAGAAAGAGTCTGGGCATG 888  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
DB 889 GAGATAACTCATCAGAAATCATCTTTTGACAGATTCTTCTGTAGTTTCAGAAATCAGACACA 948  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
DB 949 AGTTTGACAGAAAGAGAGTCTCAGCTGCTTAGATGAAGTCTTGACATTAACAGCAACATG 1008  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
DB 1009 GGCAGTGATTCAGGCACCATGGGAAGTGAATCAGATGAAGAAATGTGGCAGCAAGACA 1068  
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
DB 1069 TCCCCGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTG 1128  
QY 321 GluGlyIlyAsnIleIleIleCysLeuProThrGlySerGlyIlySerThrArgValAlaVal 340  
DB 1129 GAAGGAGAAATATCATCTCTGCTCCCTACAGGAGTGGAAAAACCGAGTGGCTGTT 1188  
QY 341 TyrIleAlaIlyAspPheIleuAspIlyIlySerGlyAlaSerGluProGlyIlyValIle 360  
DB 1189 TACATTGCCAAGATCACTTAGACAAAGAAAGAAAGCATCTGAGCCTGGAAGGTTATA 1248  
QY 361 ValLeuValAsnIlyValLeuLeuValGluGlnLeuPheArgIlyGluPheGlnProPhe 380  
DB 1249 GTTCTTGTCATAAAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAACCATTT 1308  
QY 381 LeuIlyIlySerTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuIlyIleSerPhe 400  
DB 1309 TTGAAGAAATGTATCTGCTGTTATTGATTAAAGTGATACCAACTGAATAATCATATT 1368  
QY 401 ProGluValValIlySerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
DB 1369 CCAGAACTTGTCAAGTCTCTGTGATATTATTATATCAGTACAGCTCAAAATCCTTGAACCTCC 1428  
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
DB 1429 CTCTTAAACTTGAAAAATGAGAAAGATGCTGGTGTTCAAATTGTCAAGACTTTTCCCTCATTT 1488  
QY 441 IleIleAspGlyCysHisHisThrAsnIlyGluAlaValTyrAsnAsnIleMetArgHis 460  
DB 1489 ATCATGTATGAATGTATCATCACCAACAAAGAGCAGTGTATATTAACATCATGAGGCAT 1548  
QY 461 TyrLeuMetGlnIlySerIlyAsnAsnArgLeuIlyIlyGluAsnIlyProValIlePro 480  
DB 1549 TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAACAAACCACTGATTCCC 1608  
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyIlyAlaThrIlyGlnAla 500

Db 1609 CTTCCCTGAGATACTGGAGCTAAAGCCTTCACTGGTGGAGGGGCCAGAGCAAGCC 1668  
Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
Db 1669 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACT 1728  
Qy 521 ValLysGluAsnLeuAspGluLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540  
Db 1729 GTTAAAGAAAACTTGATCAACTGAAAAACCAAAATACAGAGCCATGCAGAAAGTTTGCC 1788  
Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560  
Db 1789 ATTGCAAGATGCAACGAGAGAACTCATTTAAAGAAACTTCTAGAAATAATGACAAAG 1848  
Qy 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580  
Db 1849 ATTCAAACTTATGTGCAATGAGTCCAATGTCAAGATTTGGAACCTCAACCTATGAACAA 1908  
Qy 581 TyrAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
Db 1909 TGGGCCCATTCAAATGGAATAAAAGCTGCCAAAAAGGAAATCGCAAGAACGTGTTGT 1968  
Qy 601 AlaGluHisLeuArgLysTyrArgGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
Db 1969 GCAGAACATTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACAAATTCGAATGATA 2028  
Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640  
Db 2029 GATGCGTATCACTCATCTTGAACCTTCTATATGAAGAGAAAGATTAAGAGTTTGCAATC 2088  
Qy 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 2089 ATAGAAGATGATAGTATGAGGGGTGATGATGAGTATGTGATGAGTGAAGATGAG 2148  
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2149 GATGATTTAAAGAAACCTTGAAACTGCAATGAAACAGATGATTTCTCATGACTTATATT 2208  
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2209 TTTGAAAACAATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAAAG 2268  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720  
Db 2269 CTGACCAATTAAGAAATACCAATATGAGCAATATATACTAGCACTGAGGAATCAGACGA 2328  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGlu 740  
Db 2329 GGAATATCTTTACAAAAACAGACAGAGTGCATATGCGCTTCCAGTGCATTAAGAA 2388  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisIleLysIleGlyAlaGlyHisSer 760  
Db 2389 AATGAAAAATTTGCTGAGTAGAGTCAAAAGCCCAACCATCTGATTGGAGCTGCACACAGC 2448  
Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2449 AGTGAGTTCAAAACCATGACACAGAAATGAACAAAAAGAGTCATTAGTAAATTTGCACT 2508  
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2509 GGAAAAATCAATCTGCTTATCGCTACACAGTGCAGAGAAAGTCTGCATATTAAGAA 2568  
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2569 TGTAACATGTTATCCGTTATGCTCGTCAACCAATGAATAGCCATGATGTCAGGCCGT 2628  
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2629 GGTGAGCCGAGAGCTGATGAGAGCACTACGTCTGTTGCTCACAGTGTTCAGAGATT 2688  
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860

Db 2689 ATCGAACATGAGACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTGT 2748  
Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2749 GTTCAAAATATGAAACGAGAGAGTATGCTCATAGATTTTGAAATTACAGATGCAAAAGT 2808  
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900  
Db 2809 ATAATGAAAAAGAAATGAAAAACCAAGAGAAATATGCAAGCATTAACAAGATAACCCA 2868  
Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920  
Db 2869 TCACTAATACTTCTTTCGAAAACTGCAGTGTGCTAGCCGTCTGGGGAAGATATC 2928  
Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
Db 2929 CATGTAATTGAGAAAAATGCATCAGTCAATATATGACCCCAAGATTCAAGGAACTTTACATT 2988  
Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 2989 GTAAGAGAAAAACAAAGCACTGCAGAAAGAGTGTCCGACTATCAAAATAATGTGTGAATC 3048  
Qy 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3049 ATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAAGATTGCC 3108  
Qy 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000  
Db 3109 TGTCGAAATAAGAAATTTGTAGTGTGTTTCAAAATAATTCAACAAGAAACAATAC 3168  
Qy 1001 LysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020  
Db 3169 AAAAGTGGGTAGAATTACTATCAATTTCCCAATCTTGACTATTCAAGATGCTGTTTA 3228  
Qy 1021 PheSerAspGluAsp 1025  
Db 3229 TTTAGTGATGAGGAT 3243  
RESULT 3  
ADRI4404  
ID ADRI4404 standard, DNA; 3380 BP.  
XX  
AC ADRI4404;  
XX  
DT 21-OCT-2004 (first entry)  
XX  
DE Human NF-kappaB pathway-associated gene SeqID405.  
XX  
KW NF-kappaB pathway; antiinflammatory; cyostatic; hepatotropic; virucide;  
KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;  
KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotrophic;  
KW immunosuppressive; vulnerary; gene therapy; immune disorder;  
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;  
KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;  
KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;  
KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;  
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;  
KW viral replication; host cell survival; evasion of immune response;  
KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;  
KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAF;  
KW autoimmune disorder; hyper immune activity;  
KW aberrant acute phase response; hypercongenital condition; birth defect;  
KW necrotic lesion; wound; organ transplant rejection;  
KW aberrant signal transduction; proliferating disorder; cancer;  
KW HIV propagation; gene; ds; human.  
OS Homo sapiens.  
XX  
PN WO2004065577-A2.  
XX  
XX 05-AUG-2004.  
XX  
PD 13-JAN-2004; 2004WO-US000798.  
PF

XX 14-JAN-2003; 2003US-0440068P.  
PR 12-MAY-2003; 2003US-0469757P.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Nadler SG, Neubauer MG, Feder JN, Carman J;  
XX  
DR WPI; 2004-562168/54.  
DR P-PSDB; ADRI4405.  
XX  
PT New isolated polynucleotides and polypeptides associated with NF-kappab  
PT pathway, useful for diagnosing, treating, or preventing disorders or  
PT diseases associated with NF-kappab pathway.  
XX  
PS Claim 1; SEQ ID NO 405; 237pp; English.

CC This invention relates to the novel association of protein sequences (and  
CC the genes which encode them) to the NF-kappab pathway. The invention may  
CC be useful for the production of compounds with an antiinflammatory,  
CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,  
CC gastrointestinal-Gen, antiaesthetic, antiarteriosclerotic,  
CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or  
CC vulnerary activity or for gene therapy. The proteins and nucleotides are  
CC useful for diagnosing, preventing, treating, or ameliorating conditions  
CC or diseases associated with the NF-kappab pathway. The condition is an  
CC immune disorder, an inflammatory disorder, an inflammatory disorder  
CC related to aberrant NF-kappab regulation, cancer, aberrant apoptosis,  
CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM  
CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic  
CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,  
CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell  
CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory  
CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick  
CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper  
CC immune activity, disorders related to aberrant acute phase responses,  
CC hypercongenital conditions, birth defects, necrotic lesions, wounds,  
CC organ transplant rejection, conditions related to organ transplant  
CC rejection, disorders related to aberrant signal transduction,  
CC proliferating disorders, cancers and HIV propagation in cells infected  
CC with other viruses. The present sequence is that of a human gene which is  
CC subject to the novel association with the NF-kappab pathway of the  
CC invention. Note: This sequence does not appear in the specification but  
CC was obtained by the indexer from Genbank.  
XX  
SQ Sequence 3380 BP; 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0 Length: 3380  
Score: 5311.00 Matches: 1025  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 13 Gaps: 0

US-09-515-363C-2 (1-1025) x ADRI4404 (1-3380)

QY 1 MetSerAnGLYTrSerThraPGLuaSnPheaRGTyrLeuIleSerCysPheaRgAla 20  
DB 169 ATGTCGAATGGGTATTCCACAGACGAGAAATTCGCTATCTCATCTCGTTCAGGGCC 228  
QY 21 ArgValIysMetTyrIleGlnValGluProValIleuAspTyrLeuThrPheLeuProAla 40  
DB 229 AGGGTGAATAATGTACATCCAGGTGAGCCTGTGCTGACTAACCCTGACCTTTCGCTGCA 288  
QY 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
DB 289 GAGGTGAAGACAGATTCAAGAGCAGTCGCCACCTCCGGAAACATGACAGGCAATTGAA 348  
QY 61 LeuLeuSerThrLeuGluIysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80  
DB 349 CTGCTGCTGAGCACCTTGAGAAAGGAGTCTGGCACCTTGTTGGACTCGGAAATTCGTG 408

QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
DB 409 GAGGCCCTCCGAGAACCGGACGCCCTCTGGCGCCCGCTACATGAACCTGAGCTCACG 468  
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
DB 469 GACTTGCCCTCTCCATCGTTTGAGAACCGCTCATGATGAATATCTCCAACGCTGAACCTC 528  
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValIleuAspLysCysMetGlu 140  
DB 529 CTTCAGCCCACTCTGGTGACAAGCTTCTAGTTAGAGACGCTTGATTAAGTCATGAG 588  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAnGLYAsn 160  
DB 589 GAGGAACCTGTGACAAATTGACAGACAGAAACCGGATTGCTGCTGCAGAAACAATGGAAT 648  
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180  
DB 649 GAATCAGGTGTAGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAACCTGTTCTGCA 708  
QY 181 PheLeuAsnValIleuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
DB 709 TTTCTGAATGTTCTTCGTCAACAAGAAACAATGAACCTGTCCAGAGTTTAACAGGCTCT 768  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
DB 769 GATTGCTCAGAAAGCAATGCAGAGATTGAGAAITTAACAAAGTTGATGCTCTCAAGTG 828  
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240  
DB 829 GAAGACCAACTCTTTCACCAACACAGTTCAAGCCAAATCTGAGAAAGAGGTCTGGGCATG 888  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
DB 889 GAGAACTAATCATCAGAAATCATCTTTGACAGATTCTTGTAGTTTCAGATCAGACACA 948  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
DB 949 AGTTTGCAGAAAGAGTGTACGCTGTTAGATGAAGAATCTTGACATACAGCAACATG 1008  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
DB 1009 GGCAGTGAATTCAGGCACCATGGGAAGTGAATTCAGATGAAGAATGTGGCAGCAAGACA 1068  
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
DB 1069 TCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTG 1128  
QY 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340  
DB 1129 GAAGGAGAATATCATCATCTGCCCTCCCTACAGGAGTGGAAGAACAGAGTGCGTGT 1188  
QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle 360  
DB 1189 TACATTGCCAAGATCATCTAGACAGAGAAAAAGCATCTGAGCCTGGAAGAAATTATA 1248  
QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380  
DB 1249 GTTCTTGTCATTAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACCATTT 1308  
QY 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400  
DB 1309 TTGAAGAAATGTTATCGTTATTTGATTAAAGTGTGATACCAACTGAAAAATATCATTT 1368  
QY 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
DB 1369 CCAGAGTGTCAAGTCTGTGATATTATATCAGTACAGCTCAAAATCCTGAAAACTCC 1428  
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
DB 1429 CTCTTAACTTGAAAAATGGAAGAATGCTGTGTCAATTGTCAAGACTTTCCCTCATTT 1488  
QY 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetCArgHis 460

Db 1489 ATCATTTGATGATGTGCATCACCAACAAGAAGCAGTGATATAATACATCATGAGGCAT 1548  
Qy 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480  
Db 1549 TATTTCATGCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCACTGATTCCC 1608  
Qy 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500  
Db 1609 CTTCCTCAGATACCTGGGACTTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCC 1668  
Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
Db 1669 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACT 1728  
Qy 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540  
Db 1729 GTTAAAGAAAAACCTTGATCACTGAATAAACCAATACAGAGCCATGCAAGAGTTTGCC 1788  
Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuGluIleMetThrArg 560  
Db 1789 ATTGCAGATGCCAACCAAGAGATCCATTTAAGAGAAACTTTAGAAATATGACAAGG 1848  
Qy 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580  
Db 1849 ATTCAAACTTATTGTCAAAATGAGTCCAATGTCAGATTTTGGAACTCAACCTTATGAA 1908  
Qy 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
Db 1909 TGGGCATTCAAATGGAATAAAAGCTGCAAAAAAGGAAATCGCAAGAACTGTTTGT 1968  
Qy 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
Db 1969 GCAGACATTTGAGGAAGTACATGAGGCCCTCAACAATTATGACACAAATTCGAATGATA 2028  
Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640  
Db 2029 GATGCGTATACTCATCTTGAAACTTTCTATATGAGAGAAAGATAAGAGTTTGCACTC 2088  
Qy 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 2089 ATAGAGATGATAGTAGAGGCTGTGATGATGAGTATTGTGATGTGTGATGAGATGAG 2148  
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2149 GATGATTTAAAGAAACCTTTGAAACTGCGATGAAACAGATAGATTCTCATGACTTTAT 2208  
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2209 TTTGAAAAACAATAAATGTGAAAAAGGCTGCTGAAAAACCAAGATATGAATAAGAAAG 2268  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGlnSerAlaArg 720  
Db 2269 CTGACCAATTTAAGAAATACCAATATGAGCAATATACTAGGACTGAGGAATCAGCAGCA 2328  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740  
Db 2329 GGAATTAACTTTTACAAAAACAGACAGAGTGCATATGCGCTTCCAGTGGATTACTGAA 2388  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760  
Db 2389 AATGAAAAAATTTGCTGAAGTAGAGTCAAAAGCCCACTCTGATTGAGCTGACACAGC 2448  
Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2449 AGTGAGTTCAAAACCCATGACACAGATGACAAAAAGAGTCAATAGTAATTTCCGACT 2508  
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2509 GGAAAAATCAATCTGCTTATCGCTACACAGTGGCAGAGAGAGTCTGATATTAAAAAGAA 2568  
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820

Db 2569 TGTAACTTGTATTCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGT 2628  
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2629 GGTGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTGTCTCACAGTGGTTCAGAGATT 2688  
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860  
Db 2689 ATCGAACATGAGACAGATTATGATTTCCGAGAGAGATGATGATATAAGCTATACATTGT 2748  
Qy 861 ValGluAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2749 GTTCAAAATATGAACCAAGAGAGTATGCTCATAGATTTTGAATTACAGATGCAAGT 2808  
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900  
Db 2809 ATATGCAAAAAGAAAAATGAAAAACCAAGAGAAATATTGCCAAGCATTTACAGAAATTA 2868  
Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920  
Db 2869 TCACATAATTACTTCCCTTGCAAAAACTGCAGTGTGCTAGCCCTGTTGGGGAAGATATC 2928  
Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
Db 2929 CATGTAATTGAGAAAAATGCATCACGTCATATGACCCAGAAATTCAGAGAACTTTACATT 2988  
Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 2989 GTACAGAAAAACAAAGCACTGCACAAAGAGTGTCCGACTATCAATTAATGTGAATC 3048  
Qy 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3049 ATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGCC 3108  
Qy 981 CysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnTyr 1000  
Db 3109 TGTCTCAAAATTAAGGAATTTGTAGTGTGTTTCAAAAAATAATTCAACAAAGAAACA 3168  
Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020  
Db 3169 AAAAAGTGGGTAGAAATTACTATCACATTTCCCAATCTTGACTATTCAGAATGCTGTTA 3228  
Qy 1021 PheSerAspGluAsp 1025  
Db 3229 TTTAGTGATGAGGAT 3243  
RESULT 4  
AEA23691  
ID AEA23691 standard; DNA; 3380 BP.  
XX AEA23691;  
AC AEA23691;  
XX  
DT 11-AUG-2005 (first entry)  
XX  
DE Human PRO polypeptide DNA SEQ ID NO 233.  
XX  
KW immune disorder; PRO; Antiinflammatory; Dermatological;  
KW Immunosuppressive; Antirheumatic; Antiarthritic; Osteopathic;  
KW Muscular-Gen.; Vasotropic; Antianemic; Antipsoriatic; Immunostimulant;  
KW ds; gene.  
XX  
OS Homo sapiens.  
XX  
PN WO2005051988-A2.  
XX  
PD 09-JUN-2005.  
XX  
PF 02-MAR-2004; 2004WO-US006460.  
XX  
PR 03-MAR-2003; 2003US-0451884P.  
XX  
PA (GETH ) GENENTECH INC.  
XX

PI Abbas A, Bodary S, Clark H, Schoenfeld J, Williams PM, Wood WI;  
PI Wu TD;  
XX WPI; 2005-417958/42.  
DR P-PSDB; AEA23692.  
XX  
PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and  
PT treating an immune related disorder, e.g. systemic lupus erythematosus,  
PT rheumatoid arthritis, osteoarthritis, autoimmune hemolytic anemia, or  
PT psoriasis.  
XX  
PS Disclosure; SEQ ID NO 233; 966bp; English.  
XX  
CC The invention relates to an isolated nucleic acid. The polypeptide,  
CC compound or composition, and methods are useful for diagnosing and  
CC treating an immune related disorder, e.g. systemic lupus erythematosus,  
CC rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis,  
CC spondyloarthritis, systemic sclerosis, idiopathic inflammatory  
CC myopathies, Sjogren's syndrome, systemic vasculitis, sarcoidosis,  
CC autoimmune hemolytic anemia, autoimmune or immune-mediated skin diseases  
CC including bullous skin diseases, erythema multiforme and contact  
CC dermatitis, psoriasis, lymphadenopathy, splenomegaly and leukopenia. The  
CC present sequence represents a human PRO polypeptide DNA.  
XX  
SQ Sequence 3380 BP; 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0 Length: 3380  
Score: 5311.00 Matches: 1025  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 14 Gaps: 0

US-09-515-363C-2 (1-1025) x AEA23691 (1-3380)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheargTyrLeuIleSerCysPheArgAla 20  
DB 169 ATGTGAAATGGGTATTCACAGACGAGAAATTCGGTATCTCATCTCGTGTTCAGGGCC 228  
QY 21 ArgValIysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
DB 229 AGGGTGAATAATGTACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCGCTTCA 288  
QY 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
DB 289 GAGGTGAAGAGACAGATTCAAGAGACAGTCCGACCTCCGGGAACATGCAGCGATTGAA 348  
QY 61 LeuLeuLeuSerThrLeuGluIlySgIlyValTPrHisLeuGlyTPrThrArgGluPheVal 80  
DB 349 CTGCTGCTGACACCTTGAGAAAGGAGTCTGGCACTTGCTTGGACTCGGGAATTCTGTG 408  
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
DB 409 GAGGCCCTCCGAGAACCGGACCCCTCTGGCCGCCCTAATGAACCTGAGCTCAGC 468  
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
DB 469 GACTTGCCCTCTCCATCGTTTGAAACGCTCATGATGATATCTCCAACCTGCTGAACCTC 528  
QY 121 LeuGlnProThrLeuValAspIlySLeuLeuValArgAspValLeuAspIlySLeuMetGlu 140  
DB 529 CTTCAGCCCACTCTGTGACAGCTTCTAGTAGAGCGTCTTGATTAAGTCATGGAG 588  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaIleGluAsnAsnGlyAsn 160  
DB 589 GAGGAAGTGTGACAAATGGAAGACAGAAACCGGATTGCTGCTGCAGAAAACAATGGAAT 648  
QY 161 GluSerGlyValArgGluLeuLeuIlySArgIleValGlnIlySgIlyAsnTrpPheSerAla 180  
DB 649 GAATCAGGTGTAAAGAGAGCTACTATAAAGGATTGTGCAGAAAAGAAAAGCTGTTCTTGC 708  
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200

DB 709 TTTCTGAATGTTCTTCGTCAAAACAGAAACAAATGAACCTGTCCAAAGATTAAACAGCTCT 768  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
DB 769 GATTGCTCAGAAAGCAATGCAGAGATTGAGATTATCACAAAGTTGATGTCTCAAGTG 828  
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIlySgIlyValTPrGlyMet 240  
DB 829 GAAGAGCAACTCTTTCAACACAGATTCAAGCCAAATCTGAGAAAGAGGTCTGGGCATG 888  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
DB 889 GAGAAATTAATCATCAGAAATCATCTTTGACAGATTCTTGTAAGTTTCAGAAATCAGACACA 948  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
DB 949 AGTTTGACAGAAAGAGTGTACGCTGCTTAGTGAAGTCTTGACATTAACAGCAACATG 1008  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
DB 1009 GGCAGTGATTCAAGCACCATGGGAAGTGAATGATGAAGAAGATGTGCACAAAGACA 1068  
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
DB 1069 TCCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTG 1128  
QY 321 GluGlyIlySAsnIleIleIleCysLeuProThrGlySerGlyIlySerThrArgValAlaVal 340  
DB 1129 GAAGGGAAGATATCATCTCTGCTCCCTACAGGAGTGGAAGAAACAGAGTGGCTGT 1188  
QY 341 TyrIleAlaIlySAspHisLeuAspIlySLeuAlaSerGluProGlyIlySValIle 360  
DB 1189 TACATTGCCAAGGATCATCTTAGACAAGAGAAAAAGCATCTGAGCCTGGAAAAAGTTATA 1248  
QY 361 ValLeuValAsnIlySValLeuLeuValGluGlnLeuPheArgIlySgIlyPheGlnProPhe 380  
DB 1249 GTTCTTGTCATAAGGTACTGTAGTTGAACAGCTCTCCGCAAGAGTTCCAACCATTT 1308  
QY 381 LeuIlySLeuTPrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuIlySLeuPhe 400  
DB 1309 TTGAAGAATGTATCGTGTATTTGATTAAAGTGTGATACCAACTGMAAATATCATTT 1368  
QY 401 ProGluValValIlySLeuSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
DB 1369 CCAGAACTTGTCAAGTCCCTGTGATATTATTATATCATCAGCTCAATCCCTGAAAACTCC 1428  
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
DB 1429 CTCTTAACTTGAAAAATGAGAAAGATGTGTGTTCATTCAATTGTCAAGACTTTTCCCTCAT 1488  
QY 441 IleIleAspGluCysHisHisThrAsnIlySgIlyAlaValTyrAsnAsnIleMetArgHis 460  
DB 1489 ATCATTTGATGAATGTATCATCACCAACAAAGAGCATGTATATATCATCATGAGGCAT 1548  
QY 461 TyrLeuMetGlnIlySLeuIlySAsnAsnArgLeuIlySgIlyAsnIlySProValIlePro 480  
DB 1549 TATTGATGCAGAAAGTTGAACAAATAGACTCAAGAAAAGAAAACAAACCAAGTGAATCCC 1608  
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrIlySgIlyAla 500  
DB 1609 CTTCTCAGATACTGGGACTTAACAGCTTCACTGTGTGAGAGGGCCAGAAAGCAAGCC 1668  
QY 501 IlySAlaGluGluHisIleLeuIlySLeuCysAlaAsnLeuAspAlaPheThrIleIlySThr 520  
DB 1669 AAAGCTGAAGAACACATTTTAAACTATATGTGCCAATCTTGATGCAATTTACTATTAAACT 1728  
QY 521 ValIlySgIlyAsnLeuAspGlnLeuIlySAsnGlnIleGlnGluProCysIlySLeuPheAla 540  
DB 1729 GTTAAAGAAAACCTTGATCACTGAAAAACCAATATACAGAGCCATGCAGAAAGTTGCC 1788  
QY 541 IleAlaAspAlaThrArgGluAspProPheIlySgIlySLeuLeuGluIleMetThrArg 560

Db 1789 ATTGCAGATGCAACGAGAGATCCATTTAAGAGAACTTCTAGAATAATGACAAAGG 1848  
Qy 561 IleglnThrTyrCysGlnMetSerPrometSerAspPheGlyThrGlnProTyrGluGln 580  
Db 1849 ATTCAACTTATTTGTCMAATGAGTCCAAATGTCAGATTGTGGAATCCAACTTATGAAACA 1908  
Qy 581 TTPAlaIleglnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
Db 1909 TGGGCCATTCAAAATGAAAAAAAGCTGCAAAAAAAGGAATCGCAAGAACGTGTTGT 1968  
Qy 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
Db 1969 GCAGACATTGTGAGGAGTAGATGAGGCCCTACAAATTAATGACACAAATTCGAATGATA 2028  
Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640  
Db 2029 GATGGCTATCTCATCTTGAACCTTCTATTAATGAAGAGAAAGATAAGAGTTTGCACTC 2088  
Qy 641 IlegluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 2089 ATAGAAGATGATAGTAGAGGGTGTGATGATGAGTATGTGATGTGTGATGAAGATGAG 2148  
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2149 GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTCTCATGACTTATATT 2208  
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2209 TTTGAAAACAATAAATGTGAAAAAGGCTGGCTGAAAAACCAAGATATGAAAAAG 2268  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720  
Db 2269 CTGACCAATTTAAGAAATACCAATATGAGCAATATACTAGGACTGAGGAATCAGACGA 2328  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740  
Db 2329 GGAATTAATCTTTACAAAAACGACAGAGTGCATATGCGCTTCCAGTGGATTAAGTAA 2388  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleglyAlaGlyHisSer 760  
Db 2389 AATGAAAAATTGCTGAAGTAGAGTCAAAAGCCCAACCTGTGATTTGAGCTGACACAGC 2448  
Qy 761 SerGluPheLysPrometThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2449 AGTGAGTTCAAAACCATGACACAGAAATGAACAAAAAGAAAGTCATTAGTAAATTTCCG 2508  
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2509 GGAATAATCAATCTGCTTATCGCTACCAAGTGGCAGAGAAAGGTCTGATATTAAGAA 2568  
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2569 TGTAACATTTGTTATCCGTATGCTCTGTCACCAATGAATAGCCATGTGTCAGGCCGT 2628  
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2629 GGTGAGCCAGAGCTGATGAGACACCTACGTCTGTGCTCACAGTGGTTCAAGAGTT 2688  
Qy 841 IlegluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860  
Db 2689 ATCGACATGAGACAGTTAATGATTTCCGAGAGAGATGATGTATAAGCTATACATGTT 2748  
Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2749 GTTCAAAATATGAAACGAGAGATATGCTCATTAAGATTTTGGAAATTACAGATCGAAAGT 2808  
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisIleTyrLysAsnAsnPro 900  
Db 2809 ATAATGAAAAAGAAATGAAACCAAGAAATATTTGCCAAGCATTACAAAGATTAACCCA 2868  
Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920  
Db 2869 TCACATAATACTTTCTTGCAAAAACTGCAGTGTGTAGCCTGTTCTGGGGAAGATATC 2928

Qy 921 HisValIlegluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
Db 2929 CATGTAATTGAGAAATGCAATCACGTCAATATGACCCAGAAATTCAGAGAACTTTACATT 2988  
Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGluIle 960  
Db 2989 GTAAGAGAAAAACAAAGCACTGCCAAAAGAGTGTCCGACTATCAATAATGTGAATC 3048  
Qy 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3049 ATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGTGCACAAAGGCTTAGATTGCCCT 3108  
Qy 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000  
Db 3109 TGTCTCAAAATAAGGAATTTGTAGTGTGTTTCAAAAATAATTCAACAAAGAAACAATAC 3168  
Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLysLeu 1020  
Db 3169 AAAAAGTGGTAGAATTACTATCACATTTCCCAATCTTGACTATTGAGATGCTGTTTA 3228  
Qy 1021 PheSerAspGluAsp 1025  
Db 3229 TTTAGTGATGAGGAT 3243  
RESULT 5  
AEA36113  
ID AEA36113 standard; DNA; 3380 BP.  
XX  
AC AEA36113;  
XX  
DT 25-AUG-2005 (first entry)  
XX  
DE Human nucleic acid sequence #45.  
XX  
KW Screening; gene expression; colorectal tumor; colitis; Crohns disease;  
KW irritable bowel syndrome; gastrointestinal disease; cytostatic;  
KW gastrointestinal-gen.; antiinflammatory; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO2005054507-A2.  
XX  
PD 16-JUN-2005.  
XX  
PF 03-DEC-2004; 2004WO-GB005078.  
XX  
PR 04-DEC-2003; 2003GB-00028048.  
XX  
PA (UYSH-) UNIV SHEFFIELD.  
XX  
PI Corfe B, Chirakkal H;  
XX  
DR WPI; 2005-435407/44.  
XX  
PT Screening for nucleic acid molecules exhibiting altered expression in  
PT cells grown in the presence of butyrate, and detection of the nucleic  
PT acid molecules or the encoded polypeptides in diagnosing colorectal  
PT cancer.  
XX  
PS Disclosure; Page 77-78; 266pp; English.  
XX  
CC The invention relates to a method of screening for nucleic acid molecules  
CC that show altered expression in a first cell sample comprising comparing  
CC the gene expression profile of the sample with that of a second reference  
CC sample, where the first sample has been grown in the presence of butyrate  
CC or a related carbon source from which butyrate is directly or indirectly  
CC derived, but the reference sample has not. The invention also relates to  
CC a method of detecting at least one nucleic acid molecule associated with  
CC the initiation and/or progression of colorectal cancer in an animal,  
CC comprising providing a biological sample comprising at least one cell to  
CC be tested, contacting the sample with a ligand (preferably a hybridizing  
CC nucleic acid molecule) which binds to at least one nucleic acid and

CC detecting the presence of at least one molecule in the sample, a method  
CC of detecting at least one polypeptide associated with the initiation  
CC and/or progression of colorectal cancer in an animal comprising providing  
CC a biological sample comprising at least one cell to be tested, contacting  
CC the sample with at least one ligand that specifically binds at least one  
CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
CC an amino acid sequence which varies by the addition, deletion or  
CC substitution of at least one amino acid residue and detecting the  
CC presence of the polypeptide in the sample, a method of screening for  
CC agents that modulate the activity of at least one polypeptide encoded by  
CC a gene associated with the initiation and/or progression of colorectal  
CC cancer comprising forming a preparation comprising at least one  
CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising  
CC at least one addition, deletion or substitution and at least one agent to  
CC be tested and determining the activity of the agent with respect to  
CC activity of the polypeptide, and an antibody or its effective binding  
CC portion identified by the method, for use as a pharmaceutical. The  
CC methods are useful for screening for nucleic acid molecules that show  
CC altered expression in a cell sample, and for detecting a nucleic acid and  
CC a polypeptide respectively, that are associated with the initiation  
CC and/or progression of colorectal cancer and are useful for detecting or  
CC monitoring colorectal cancer, especially adenocarcinoma. The methods are  
CC also useful for screening for agents that modulate the activity of at  
CC least one polypeptide encoded by a gene associated with the initiation  
CC and/or progression of cancer, where agents identified by the method are  
CC useful for treating colorectal cancer. The methods could also be used to  
CC detect or monitor other conditions such as colitis, Crohn's disease or  
CC irritable bowel syndrome, as a screening tool for fiber consumption, as  
CC an assay for colon microflora functionality or for early detection of pre  
CC -cancerous growth. This sequence represents a human nucleic acid  
CC identified by the screening method of the invention.

XX  
SQ Sequence 3380 BP; 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0 length: 3380  
Score: 5311.00 Matches: 1025  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 14 Gaps: 0

US-09-515-363C-2 (1-1025) x AEA36113 (1-3380)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
DB 169 ATGTCGAATGGGTATTCACAGACGAGAATTTCCGCTATCTCATCTCGTGCTTCAGGGCC 228  
QY 21 ArgValIleMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
DB 229 AGGGTGAATAATGATCATCCAGGTGAGCCTGTGCTGACTACCTGACCTTTCGCTGCA 288  
QY 41 GluValIleGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
DB 289 GAGGTGAAGAGCAGATTCAAGAGCACTCGCCACCTCCGGGAACATGCAAGGCAGTTGAA 348  
QY 61 LeuLeuLeuSerThrLeuGluIleGlyValTyrPheIleGlyTyrThrArgGluPheVal 80  
DB 349 CTGCTGCTGACACCTTGAGAGAGGAGTCTGGCACCTGGTTGGACTCGGGAATTGCTG 408  
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
DB 409 GAGGCCCTCCGGAGAACCGGAGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCAGC 468  
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
DB 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACCTC 528  
QY 121 LeuGlnProThrLeuValAspIleLeuLeuValArgAspValLeuAspIleCysMetGlu 140  
DB 529 CTTCAAGCCCACTCTGGTGACAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCAATGGAG 588  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160

DB 589 GAGGAAGCTGTGACAAATGGAAGACAGAAACCGGATTTGCTGCTGCAGAAACAAATGGAAT 648  
QY 161 GluSerGlyValArgGluLeuLeuIleValGlnIleGlyAsnIlePheSerAla 180  
DB 649 GAATCAGGTGTAAAGAGCTAACTAAAAAGGATTTGCAGAAAAGAAAACCTGTTCTTGCA 708  
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
DB 709 TTTCTGAATGTTCTTCGTCAACAGGAAACAATGAATTTGCCAAGAGTTAAACAGCTCT 768  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
DB 769 GATTGCTCAGAAAGCAATGCAGAGATTGAGATTTATCACAAGTTGATGGTCTCAAGTG 828  
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIleGluValTyrGlyMet 240  
DB 829 GAAGAGCAACTTCTTCAACACACAGTTCAAGCCAAATCTGGAGAAAGAGGCTGGGGCATG 888  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
DB 889 GAGAAATACTCATCAGAAATCATCTTTGCAGATTCTCTGTAGTTTCAAGATCAGACACA 948  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
DB 949 AGTTGGCAGAGAGAGAGTGCAGCTGCTTAGATGAAGTCTTGACATTAACAGCAACATG 1008  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
DB 1009 GGCAGTGATTCAGGCACCATGGGAAGTGATTCAAGTGAAGAGATGTGGCAGCAAGACA 1068  
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
DB 1069 TCCCCGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAAGTTGCCCGCAGCCCTTG 1128  
QY 321 GluGlyIleAsnIleIleIleCysLeuProThrGlySerGlyIleThrArgValAlaVal 340  
DB 1129 GAAGGAGAAATATCATCATCTGCCCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGT 1188  
QY 341 TyrIleAlaIleAspHisIleAspIleValIleValIleValIleValIleValIleVal 360  
DB 1189 TACATTCGCAAGAGATCATTAGACAAAGAAAAAGCATCTGAGCCTGGAAGTATATA 1248  
QY 361 ValLeuValAsnIleValLeuValGluGlnLeuPheArgIleGluPheGlnProPhe 380  
DB 1249 GTTCTTGCAATAAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAAACATTT 1308  
QY 381 LeuIleValIleTyrArgValIleGlyLeuSerGlyAspThrGlnLeuIleIleSerPhe 400  
DB 1309 TTGAAGAAATGCTATCGTGTATTGATTAAAGTGGTGAATACCAACTGAATAATCATTT 1368  
QY 401 ProGluValValIleValSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
DB 1369 CCAGAAAGTTGTCAAGTCTGTGATATTATTATCAGTACAGCTCAAAATCTTGAAAACTCC 1428  
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
DB 1429 CTCTTAACTTGAAAAATGGAAGATGCTGGTGTTCATTTGTCAGACTTTCCTCATTT 1488  
QY 441 IleIleAspGluCysHisIleThrAsnIleGluAlaValTyrAsnAsnIleMetArgHis 460  
DB 1489 ATCATGTGATGATGTATCATCACCAACAAGAGCAGTGTATAATAACATCATGAGGCAT 1548  
QY 461 TyrLeuMetGlnIleValLeuValAsnAsnArgLeuLeuValGluAsnIleProValIlePro 480  
DB 1549 TATTTGATGACAGAGTTGAATAACAATAGACTCAAGAAAGAAACAACAGTGATTTCCC 1608  
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrIleGlnAla 500  
DB 1609 CTTCTCAGATTACTGGAGTAAACAGCTTCACTGGTGTGGAGGGGCCACGAGCAAGGCC 1668  
QY 501 LysAlaGluGluHisIleLeuValLeuCysAlaAsnLeuAspAlaPheThrIleIleYThr 520

Db 1669 AAAGCTGAAGAACACATTTTAAACTATATGTGCCAATCTTGATGCATTACTATTAAAACT 1728  
Qy 521 VallysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540  
Db 1729 GTTAAAGAAAACCTTGATCAATGAAAAACCAATATACAGAGCCATGCAGAACTTGGCC 1788  
Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuGluIleMetThrArg 560  
Db 1789 ATTGCAGATGCACACGAGAGATCCATTTAAGAGAACTTCTAGAATAATGACAAGG 1848  
Qy 561 IleGlnThrTyrCysGlnMetSerPrometSerAspPheGlyThrGlnProTyrGluGln 580  
Db 1849 ATTCAAACTTATTGTCAAAATGAGTCCAAATGTCAAGATTTGGAACTCAACCCCTATGAAACA 1908  
Qy 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
Db 1909 TGGGCATTCAAAATGAAAAAAAGCTGCAAAAAAAGAAATCGCAAGAACGTGTTGT 1968  
Qy 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
Db 1969 GCAGACATTTGAGGAAGTACATGAGGCCCTCACAAAATTAATGACACAAATTCGAATGATA 2028  
Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysAspLysLysPheAlaVal 640  
Db 2029 GATGCGTATACTCATCTTGAACCTTCTATAATGAGAGAAAGATAAGAGTTTGCAAGTC 2088  
Qy 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 2089 ATAGACATGATAGTAGGAGGTGCTGATGATGAGTATTGTGATGCTGATGATGAGATGAG 2148  
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2149 GATGATTTAAAGAAACCTTTGAAACTGATGAAACAGATAGATTCTCATGACTTTATT 2208  
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2209 TTTGAAAACAATAAAATGTTGAAAAAGGCTGGCTGAAAAACCAAGATATGAAAAAG 2268  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGlnSerAlaArg 720  
Db 2269 CTGACCAATTAAGAAATACCAATAATGAGCAATATATAGAGACTGAGAACTGACACGA 2328  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740  
Db 2329 GGAATTAATCTTACAAAAACAGACAGAGTGCATATGCGCTTCCAGTGCATTACTGAA 2388  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760  
Db 2389 AATGAAAAATTGCTGAGTAGAGTCAAAAGCCACCATCTGATTGGAGCTGACACAGC 2448  
Qy 761 SerGluPheLysPrometThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2449 AGTGAGTTCAAAACCATGACACAGATGAACAAAAAGAGTCATTAGTAAATTTCGCACT 2508  
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2509 GGAAAAATCAATCTGCTTATCGCTACACAGTGGCAGAAAGGCTCGATATTAAAAAGAA 2568  
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2569 TGTAACATGTTATCCGTTATGCTCTCGTACCAATGAATAAGCCATGCTCCAGGCCGT 2628  
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2629 GGTGAGCCAGAGCTGATGAGGACCTACGTCCTGTTGCTCACAGTGGTTCAAGAGATT 2688  
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860  
Db 2689 ATGCAACATGAGACAGTTAATGATTCCGAGAGAAAGATGATTAAGCTAATACATTGT 2748  
Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2749 GTTCAAAATATGAAACAGAGAGATATGCTCATAAGATTTTGAATTACAGATGCAAAAGT 2808

Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900  
Db 2809 ATATGGAAGAAAAGAAATGAAAAACCAAGAGAAATATTGCCAAGCATTACAGAAATAACCA 2868  
Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920  
Db 2869 TCATTAAATTACTTCTTCGCAAAAACTGCAGTGTCTAGCCTGTTCTGGGAGAGATATC 2928  
Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
Db 2929 CATGTAATTGAGAAAATGCAATCAGCTCAATATGACCCAGAAATTCAGGAACCTTACATT 2988  
Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 2989 GTAGAGAAAAACAAGCACTGCACAAAAGAGTGTGCCGACTATCAAAATTAATGCTGAATC 3048  
Qy 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3049 ATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTCACAAAGCCTTAGATTGCCCT 3108  
Qy 981 CysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnTyr 1000  
Db 3109 TGTCTCAAAATTAAGGAATTTTGTAGTGTGTTTCAAAAATTAATTCACAAACAAACAATAC 3168  
Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020  
Db 3169 AAAAAGTGGTAGAATTAATTAATCATATTCCTCAATCTTGACTATTGAGATGCTGTTA 3228  
Qy 1021 PheSerAspGluAsp 1025  
Db 3229 TTTAGTGATGAGGAT 3243  
  
RESULT 6  
ADY17563  
ID ADY17563 standard; DNA; 3434 BP.  
XX  
AC ADY17563;  
DT 05-MAY-2005 (first entry)  
XX  
DE DNA encoding a PRO polypeptide, SEQ ID NO 3369.  
XX  
KW Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;  
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;  
KW Antichryoid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;  
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;  
KW Antiallergic; ds; gene; diagnosis.  
XX  
OS Homo sapiens.  
XX  
PN WO2005016962-A2.  
XX  
PD 24-FEB-2005.  
XX  
PF 11-AUG-2004; 2004WO-US026249.  
XX  
PR 11-AUG-2003; 2003US-0493546P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;  
XX  
DR WPI; 2005-182330/19.  
XX  
PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and  
PT treating an immune related disorder, e.g. systemic lupus erythematosus,  
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.  
XX  
PS Claim 1; SEQ ID NO 3369; 158pp; English.  
XX  
CC The invention relates to an isolated nucleic acid encoding a PRO  
CC polypeptide. The polypeptide, agonist or an antagonist, antibody,

CC composition, and method are useful for diagnosing and treating an immune  
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid  
CC arthritis. The present sequence represents a PRO polypeptide.

XX Sequence 3434 BP; 1157 A; 673 C; 768 G; 836 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0 Length: 3434  
Score: 5307.00 Matches: 1024  
Percent Similarity: 100.0% Conservative: 1  
Best Local Similarity: 99.9% Mismatches: 0  
Query Match: 99.9% Indels: 0  
DB: 14 Gaps: 0

US-09-515-363C-2 (1-1025) x ADY17563 (1-3434)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
DB 223 ATGTCGAATGGGTATTCACAGACGAGAATTCCGCTATCTCATCTCGTGTTCAGGGCC 282  
QY 21 ArgValIlyMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
DB 283 AGGGTGAATAATGTACATCCAGGTGAGCCTGTGCTGGACTAACCCTGACTTTCGCTTCA 342  
QY 41 GluValIlySGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
DB 343 GAGGTGAAGAGACAGATTCAAGAGACAGTCCGACCTCCGGAAACATGCAGCAGTTGAA 402  
QY 61 LeuLeuLeuSerThrLeuGluIlySGlyValTyrPheIleGluGlyTyrThrArgGluPheVal 80  
DB 403 CTGCTGCTGAGCACCTTGAGAAAGGAGTCTGCACCTTGTTGACTCGGGAATTCCGTG 462  
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
DB 463 GAGGCCCTCCGGAGAACCGGCCCTCTGGCCGCCCTACATGAACCCGTGAGCTCAGC 522  
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
DB 523 GACTTGCCCTCTCCATCGTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 582  
QY 121 LeuGlnProThrLeuValAspIlyLeuLeuValArgAspValLeuAspIlyCysMetGlu 140  
DB 583 CTTCAGCCCACTCTGTGAGCAAGCTTCTAGTAGAGCGTCTTGATTAAGTCATGAG 642  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaIleGluAsnAsnGlyAsn 160  
DB 643 GAGGAACGTGTGACAATTGAAGACAGAAAACGAGTTGCTGCTGCAGAAAACAATGAAAT 702  
QY 161 GluSerGlyValArgGluLeuLeuIlyAsnArgIleValGlnIlySGluAsnTrpPheSerAla 180  
DB 703 GAATCAGGTGTAAAGAGAGCTACTAAAGAGATTGTGCAGAAAAGAACTGGTCTCTGCA 762  
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
DB 763 TTTTCTGAATGTTCTTCGTCAACAAGAAAACAATGAACCTTGCCAAAGTTAAACAGGCTCT 822  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
DB 823 GATTGCTCAGAAAGCAATGCAGAGATTGAAGATTATCAAGTTGATGGTCTCTCAAGTG 882  
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIlySGluValTyrGlyMet 240  
DB 883 GAAGAGCAACTTCTTCAACCAAGTTCAGCCAAATCTGAGAAAGAGGTCTGGGGCATG 942  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValIleSerGluSerAspThr 260  
DB 943 GAGAAATACTCATCAGATCATCTTTGCAGATTCTTGTAGTTTCAGAAATCAGACACA 1002  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
DB 1003 AGTTTGGCAGAAAGAGAGTGTCAAGTCTTATGATGAAAAGTCTTGACATTAACAGCACATG 1062  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300

DB 1063 GGCAGTGAATCAGGCACCATGGGAAGTGAATCAGATGAAGAGAATGTGCAGCAAGAGCA 1122  
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
DB 1123 TCCCGGAGCCAGAACTCCAGCTCAGCGCTTACCAATGGAAGTTGCCAGCCCTTG 1182  
QY 321 GluGlyIlyAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340  
DB 1183 GAAGGAAGAATATCATCTCTGCCCTCCCTACAGGAGTGGAAGAAACAGAGTGCTGTT 1242  
QY 341 TyrIleAlaIlyAspHisLeuAspIlySlySlyAlaSerGluProGlyIlyValIle 360  
DB 1243 TACATTGCCAAGGATCACTTAGACAAAGAAAGAAAGCATCTGAGCCTGCAAAAGTTATA 1302  
QY 361 ValLeuValAsnIlySlyValLeuLeuValGluGlnLeuPheArgIlySGluPheGlnProPhe 380  
DB 1303 GTTCTTGCAATAAAGTACTGCTAGTTGAACAGCTCTTCGCAAGAGTTCCAACCATTT 1362  
QY 381 LeuIlySlySTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuIlySlySerPhe 400  
DB 1363 TTGAGAAATGTATCGTGTATTGATTAGTGGTGATACCCAACTGAATAATATCATTT 1422  
QY 401 ProGluValValIlySerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
DB 1423 CCAGAGTTGTCAAGTCCGTGATATTATATCAGTACAGCTCAAAATCCTTGAAAACTCC 1482  
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
DB 1483 CTCTTAAACTTGAAAATGGAAGAAGATGCTGGTTCATTTGTCAGACTTTCCTCATTT 1542  
QY 441 IleIleAspGluCysHisHisThrAsnIlySGluAlaValTyrAsnAsnIleMetArgHis 460  
DB 1543 ATCATTTGATGAATGTATCAACCAACAAAGAGAGTGTATATATACATCATGAGGCAT 1602  
QY 461 TyrLeuMetGlnIlySlyLeuIlySlyAsnAsnArgLeuIlySlySGluAsnIlySlyProValIlePro 480  
DB 1603 TATTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAACAACCAAGTATCCC 1662  
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrIlySGlnAla 500  
DB 1663 CTTCCTCAGATACTGGGACTAAACAGCTTCACCTGTGTGAGAGGGCCAGAAAGCAAGCC 1722  
QY 501 LysAlaGluGluHisIleLeuIlySlyLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
DB 1723 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAGCT 1782  
QY 521 ValIlySGluAsnLeuAspGlnLeuIlySlyAsnGlnIleGlnGluProCysIlySlyPheAla 540  
DB 1783 GTTAAAGAAAACCTTGATCAACTGAAAAACCAATAACAGAGCCATGCAGAAAGTTGGCC 1842  
QY 541 IleAlaAspAlaThrArgGluAspProPheIlySGluIlySlyLeuLeuGluIleMetThrArg 560  
DB 1843 ATTGCAGATGCAAACCAAGAGATCCATTTAAAGAGAACTTCTAGAAATTAATGACAAAG 1902  
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580  
DB 1903 ATTCAACTTATTGTCAATGAGTCCAAATGTCAAGTTTGGAACTCAACCTATGAAACA 1962  
QY 581 TyrAlaIleGlnMetGluIlySlySlyAlaAlaIlySlySGlyAsnArgIlySGluArgValCys 600  
DB 1963 TGGGCCATTCAATGGAAGAAAGAAAGCTGCAGAAAGAAAGAAATCGCAAGAAAGTGTGT 2022  
QY 601 AlaGluHisLeuArgIlyTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
DB 2023 GCAGAACATTTGAGGAGTACATGAGGCCCTACAAATTAATGACACAAATTCGAATGATA 2082  
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluIlySlyAspIlySlyPheAlaVal 640  
DB 2083 GATGCGTATACTCATCTTGAAGCTTTCTATATATGAAGAAAGATTAAGAGTTTGCAAGTC 2142  
QY 641 IleGluAspAspSerAspGluGlyIlyAspAspGluTyrCysAspGlyAspGluAspGlu 660

Db 2143 ATAGAGATGATAGTATGAGGGTGGTATGATGAGTATTGTGATGGTGATGAAGATGAG 2202
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
Db 2203 GATGATTTAAAGAAACCTTTGAAACCTGGATGAAACAGATGATTTCTCATGACTTTATTT 2262
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
Db 2263 TTTGAAAAACAATAAATGTTGAAAAAGGCTGGCTGAAAAACCAAGATATGAAAAAG 2322
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
Db 2323 CTGACCAATTAAGAATACCAATAATGAGCAATATACTAGACTGAGGAATCAGCAGCA 2382
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
Db 2383 GGAATAATCTTTACAAAAACAGACAGAGTGCATATAGCGCTTCCAGTGGATTACTGAA 2442
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer 760
Db 2443 AATGAAAAATTGCTGAAGTAGAGTCAAGCCACCACTCTGATTGGAGCTGACACAGC 2502
Qy 761 SerGluPheLysPrometThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
Db 2503 AGTGAGTCCAACCACCATGACACAGATGACAAAAAGAGTCATTAGTAATTTCCGACT 2562
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
Db 2563 GGAAAAATAAATCTGCTTATCGCTACACAGTGGCAGAGAAGGTCCTGATATTAAGAA 2622
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db 2623 TGTAACATTGTTATCCGTTATGGTCTCGTACCAATGAATAGCCATGGTCCAGGCCCGT 2682
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
Db 2683 GGTCGAGCCAGAGCATGATGAGACACCTACGTCCTGGTCTCACAGTGGTTCAGAGATT 2742
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
Db 2743 ATCGAACATGAGACAGTTAATGATTTCCGAGAGAAAGATGATATAAGCTATACATTGT 2802
Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
Db 2803 GTTCAAATATGAACACGAGAGATATGCTCATAGATTTTGGATTAACAGATCAAAGT 2862
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
Db 2863 ATATGGAAGAAAGAAATGAAACCAAGAAATATGCCAAGCATTAACAAGATTAACCCA 2922
Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db 2923 TCACTAATACTTCTTTCGCAAAAACTGACAGTGTGCTAGCCTGTTCTGGGGAAGATATC 2982
Qy 921 HisValIleGluLysMetHisIleValAsnMetThrProGluPheLysGluLeuTyrIle 940
Db 2983 CATGTAATTGAGAAATGATCATGATATATGACCCAGAAATCAAGGAATTTACATT 3042
Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 3043 GTAAGAGAAAAACAAGCACTGCAAAAGAAGTGTGCCGACTATCAATTAATGTGAATC 3102
Qy 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db 3103 ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGGCCT 3162
Qy 981 CysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db 3163 TGCTCAAAATTAAGGAATTTGTAGTGGTTTCAAAAAATAATTCAACAAAGAAACAATAC 3222
Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020
Db 3223 AAAAAGTGGTAGAATTAACCTATCACATTTCCCAATCTTGACTATTCAAGATGCTGTTTA 3282

Qy 1021 PheSerAspGluAsp 1025
Db 3283 TTTAGTGATGAGAT 3297
RESULT 7
ADY20376
ID ADY20376 standard; DNA; 3434 BP.
XX
AC ADY20376;
DT 05-MAY-2005 (first entry)
XX
DE DNA encoding a PRO polypeptide, SEQ ID NO 6182.
XX
KW Antinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW Antichyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotropic;
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW Antiallergic; ds; gene; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2005016962-A2.
XX
PD 24-FEB-2005.
XX
PF 11-AUG-2004; 2004WO-US026249.
XX
PR 11-AUG-2003; 2003US-0493546P.
XX
PA (GETH ) GENENTECH INC.
XX
PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
XX
DR WPI; 2005-182330/19.
XX
PT New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating an immune related disorder, e.g. systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
PS Claim 1; SEQ ID NO 6182; 158bp; English.
XX
CC The invention relates to an isolated nucleic acid encoding a PRO
CC polypeptide. The polypeptide, agonist or an antagonist, antibody,
CC composition, and method are useful for diagnosing and treating an immune
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC arthritis. The present sequence represents a PRO polypeptide.
XX
SQ Sequence 3434 BP; 1157 A; 673 C; 768 G; 836 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 3434
Score: 5307.00 Matches: 1024
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.9% Mismatches: 0
Query Match: 99.9% Indels: 0
DB: 14 Gaps: 0
US-09-515-363C-2 (1-1025) x ADY20376 (1-3434)
Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
Db 223 ATGTGGAATGGGTATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTCTTCAAGGCC 282
Qy 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
Db 283 AGGGTGAATAATGTACATCAGGTGAGCGCTGTGGACTACCTGACCTTTCTGCGTGA 342
Qy 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
Db 343 GAGGTGAAGAGACAGATTACAGAGACAGTCCGCACTCCGGGAACATGACGCGAGTTGA 402

QY	61	LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal	80
Db	403	CTGCTGCTGAGCACTTGAGAAAGGAGTCTGGCACTTGTTGACTCGGAATTCTGTG	462
QY	81	GluAlaLeuArgArgThrGlySerProLeuAlaIaArgTyrMetAsnProGluLeuThr	100
Db	463	GAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCCGCCCTACATGAACCTGAGCTCACG	522
QY	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	523	GACTTGCCCTCTCCATCTGTTGAGAACGCTCATGATGAATATCTCCAATCTGTGAACCTC	582
QY	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
Db	583	CTTCAGCCCACTCTGGTGAGCAAGCTTCTAGTTAGACAGCTCTTGATTAAGTCATGGAG	642
QY	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	643	GAGGAAGTGTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAAACAATGGAAT	702
QY	161	GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTyrPheSerAla	180
Db	703	GAATCAGGTGTAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGGTTCTTGCA	762
QY	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	763	TTTCTGAATGTTCTTCGCAAAACAGAAAAACAATGAATCTGTCCAAGATTAAAGGCTCT	822
QY	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGluValAspGlyProGlnVal	220
Db	823	GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTATCACAAAGTTGATGTCCTCAAGTG	882
QY	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240
Db	883	GAAAGCAACTTCTTCAACCAACAGTTCAGCCAAATCTGGAGAAAGAGTCTGGGGCATG	942
QY	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260
Db	943	GAGAATACTCATCAGAAATCATCTTTGACAGATTCTTCTGATGTTCAGAATCAGACACA	1002
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
Db	1003	AGTTTGGCAGAGAGAGTGTCACTGCTTAGATGAAGCTTGGACATTAACAGCAACATG	1062
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	1063	GGCAGTGATTCAGGCACCATGGGAAGTGATTCAGATGAAGAAATGTGGCAGCAGACACA	1122
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
Db	1123	TCCCCGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAAGTTGCCAGCCAGCCTTG	1182
QY	321	GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
Db	1183	GAAAGGAAGATATCATCTCTGCCCTCCCTACAGGAGTGGAAAAAACAGAGTGGCTGTT	1242
QY	341	TyrIleAlaIaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle	360
Db	1243	TACATTGCCAAGGATCATTAGACAAAGAAAAAGCATCTGAGCCCTGAAAAAGTTATA	1302
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380
Db	1303	GTTCTTGTCAATAAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACCAATTT	1362
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1363	TTGAAGAAATGATCTGTATTGATTAAAGTGTGATACCCAACCTGAAAAATATCATTTT	1422
QY	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1423	CCAGAAAGTTGTCAAGTCTCTGTATATTATTATCATCAGTACAGTCCAATCCTTGAAAACTCC	1482
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440

Db	1483	CTCTTAAACTTGAAAATGGAGAGATGCTGCTGTTCAATTGTCAAGACTTTCCCTCATTT	1542
Qy	441	ILEILEASPGIUCYSHISHISTHRAASNLYSGIUALAVALTYRAASNILEMETARGHIS	460
Db	1543	ATCATTTGATGATGTATCATCACCAACAAAGAGAGGTATATATTAACATCATGAGGCAT	1602
Qy	461	TYRLEUMETGLNLYSLEULYSASNANARGLEULYSLYSGIUAASNLYSPROVALILEPRO	480
Db	1603	TATTTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAAGAAAAACAACCACTGATTCCC	1662
Qy	481	LEUPROGINILELEUGLYLEUTHRALASERPROGLYVALIGLYALATTHRYSGLNALA	500
Db	1663	CTTCCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGCCACGAAGCAAGCC	1722
Qy	501	LYSALAGIUGIUIHSILEULYSLEUCYSALASNLEUASPALAPHETHRILELYSTHR	520
Db	1723	AAAGCTGAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTACTATTAAAACT	1782
Qy	521	VALLESGIUAASNLEUASPGIUNLEULYSANGLNILEGINLUPROCYSLYSIPHEALA	540
Db	1783	GTTAAAGAAAACCTTGATCATCACTGAAAAACCAATACAGAGCCATGCAAGAAGTTTGCC	1842
Qy	541	ILEALASPALATHRRARGIUAASPPOPHELYSGIULYSLEULEUGIUILEMETTHRARG	560
Db	1843	ATTGCAGATGCACACAGAGAGATCCATTTAAAGAGAACTTCTAGAATAATGACAAG	1902
Qy	561	IILEGINTHRYCYSGIIMETSERPROMETSERASPPHEGLYTHRGINPROTYRGUGLN	580
Db	1903	ATTCAAACTTATTTGTCMAATGAGTCCAAATGTCAAGATTTGGAACTCAACCTTGAACAA	1962
Qy	581	TRPALALEGINMETGIULYSLYSALALALYSLSGLYASNARGLYSGIUAARGVALCYS	600
Db	1963	TGGGCCATTCAAATGGAATAAAAGCTGCAAAAGAAAGAAATCGCAAGAACGTGTTGT	2022
Qy	601	ALAGIUIHSLEUARGLYSTYRASNGIUALALEUGINILEASNAPTHRILEARGMETILE	620
Db	2023	GCAGACATTTGAGGAAGTACACATGAGGCCCTACAAATTAATGACACAATTCGATGATA	2082
Qy	621	ASPALATYRTHRHSILEUGIUTHRRHETTYRASNGIUGIULYSASPLYSLYPHEALAVAL	640
Db	2083	GATGCGTATCTCATCTTGAACCTTTCTATTAATGAAGAGAAAGATTAGAGTTTGCAATC	2142
Qy	641	IILEGUASPAASPSERASPGIUGLYGIYASPAASGLUTYRCYASASPGIYASPGIUAASPGIU	660
Db	2143	ATAGAAGATGATAGTATGATGAGGCTGATGATGATGATGATGATGATGATGATGATGATG	2202
Qy	661	ASPAASPLEULYSLEUASPROLEULYSLEUASPGIUTHRASPARGPHELEUMETTHREUPHE	680
Db	2203	GATGATTTAAAGAAACCTTTGAACTGGATGAAACAGATAGATTCTCATGACTTAAATTT	2262
Qy	681	PHEGIUASNAASNLYSMETLEULYSARGLEUALAGIUAASNPROGLUTYRGIUAASNGIULYS	700
Db	2263	TTTGAAAACAATAAAATGTTGAAAAGCGCTGCTGAAAACCCAGAAATATGAAAATGAAAAAG	2322
Qy	701	LEUTHRILYSLEUARGASNTHRILEMETGIUGINTYRTHRARGTHRGIUGIUSERALARG	720
Db	2323	CTGACCAAAATTAAGAAATACCATTAATGAGCAATATACTAGGACTGAGGAATCAGCAGA	2382
Qy	721	GLYILEILEPETHRILYSTHRARGIUSERALATYRALAUSERGINTRPILETHRGIU	740
Db	2383	GGAATAATCTTACAAAAACACGACAGAGTGACATATGCGCTTCCCAAGTGATTACTGAA	2442
Qy	741	ASNGIULYSIPHEALAGIUALGIVALLYVALYSAALAHISHILEULILEGLIYALAGIYHISER	760
Db	2443	AATGAAAAATTGCTGAAGTAGAGATCAAAAGCCCAACCATCTGATTGGAGCTGACACAGC	2502
Qy	761	SERGIUPHELYSPROMETTHRGINASNGIUGIUNLYSGIUALIILESERLYSPHEARGTHR	780
Db	2503	AGTGAGTTCAAAACCATGACACAGAAATGAACAAAAAGAGATCAATGATAATTTCCGACT	2562
Qy	781	GLYLYSILEASNLEULILEALATHRRVALAGIUGIUGIYLEUASPILEYSGIU	800

Db 2563 GGAAAAATAATCTGCTTATCGCTACCAACAGTGGCAGAAGAAGGCTTGATATTAAAGAA 2622  
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2623 TGTACATTTGTTATCCGTTATGCTCTCGTACCAATGAATAAGCCATGCTCAGGCCGCT 2682  
QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2683 GGTGAGCCAGAGCTGATGAGACACCTACGCTCGTGTGCTCACAGTGGTTCAAGAGATT 2742  
QY 841 ILeGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860  
Db 2743 ATCGAACATGACAGACGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTGT 2802  
QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2803 GTTCAAAATATGAAACCAAGAGAGTATGCTCATTAAGATTTTGAATTACAGATGCAAAAGT 2862  
QY 881 ILeMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900  
Db 2863 ATATGCAAAAAGAAATGAAAAACAAGAAATATTGCCAAGCATTACAAAGATAAACCCA 2922  
QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920  
Db 2923 TCACTAATAACTTCTTGGCAAAAACGTGAGTGTGCTAGCCCTGTTCTGGGGAAGATATC 2982  
QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
Db 2983 CATGTAATTGAGAAATGCATTCACGTCAATATGACCCAGAATTCAAGGAACCTTTACATT 3042  
QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 3043 GTAAAGAGAAAAACAAGCACCTGCAAAAAGAAAGTGTGCCGACTATCAATAAATGCGTGAATC 3102  
QY 961 ILeCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3103 ATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGGCCT 3162  
QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000  
Db 3163 TGTCTCAAAATAAGGAATTTGTAGTGGTTTCAAAATAATAATTCAACAAGAAACAATAC 3222  
QY 1001 LysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGlyCysLeu 1020  
Db 3223 AAAAAGTGGTAGAATTACCTATCACTTTCCCAATCTTGACTATTCAGAATGCTGTTTA 3282  
QY 1021 PheSerAspGluAsp 1025  
Db 3283 TTTAGTGATGAGGAT 3297  
RESULT 8  
ACN91892  
ID ACN91892 standard; DNA; 3928 BP.  
XX AC ACN91892;  
XX 02-DEC-2004 (first entry)  
XX DT  
XX DE Breast cancer related marker, seq id 13042.  
XX DE  
XX KW Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.  
XX OS Homo sapiens.  
XX PN US2003099974-A1.  
XX PD 29-MAY-2003.  
XX PF 18-JUL-2002; 2002US-00198846.  
XX PR 18-JUL-2001; 2001US-0306220P.  
XX PA (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Xu Y, Wang Y, Steimann K;  
XX WPI; 2003-787014/74.  
XX Novel isolated polypeptide associated with breast cancer, useful for  
PT detecting presence of polypeptide in sample, as a marker for breast  
PT cancer.  
XX PS Disclosure; SEQ ID NO 13042; 36pp; English.  
XX The invention relates to an isolated polypeptide (I) associated with  
CC breast cancer which is encoded by a nucleic acid molecule comprising a  
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to  
CC the polypeptide of the invention. The activity of the polypeptide of the  
CC invention may be described as cytostatic. The antibody is useful for  
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the  
CC invention are useful in the detection of breast tumours. (I) is useful as  
CC a marker for breast cancer and in breast cancer therapy. Sequences given  
CC in records ACN78851-ACN92934 represent nucleic acid markers associated  
CC with breast cancer. Note: The sequence listing does not form part of the  
CC specification but may be obtained in electronic format from the USPTO web  
CC site at [seqdata.uspto.gov/sequence.html?docID=20030099974](http://seqdata.uspto.gov/sequence.html?docID=20030099974)  
XX SQ Sequence 3928 BP; 1285 A; 826 C; 895 G; 917 T; 0 U; 5 Other;  
Alignment Scores:  
Pred. No.: 0 Length: 3928  
Score: 5303.00 Matches: 1023  
Percent Similarity: 99.9% Conservative: 1  
Best Local Similarity: 99.8% Mismatches: 1  
Query Match: 99.8% Indels: 0  
DB: 11 Gaps: 0  
US-09-515-363C-2 (1-1025) x ACN91892 (1-3928)  
QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
Db 611 ATGTGCAATGGGTATTCACAGACGAGAATTTCCGCTATCTCATCTGCTTCAGGGCC 670  
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
Db 671 AGGGTGAATAATGTACATCCAGGTGAGCCTGTGCTGGACTACCTGACCTTCTGCTGCA 730  
QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 731 GAGGTGAAGAGACAGATTTCAGAGGACAGTCCGCACTCCGGGACATGACGCACTGAA 790  
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrHisLeuGlyTyrThrArgGluPheVal 80  
Db 791 CTGCTGCTGAGCACCTTGAGAGAAAGGAGTCTGGCACTTGTTGACTCGGGAATTCTGTG 850  
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
Db 851 GAGGCCCTCCGAGAACCGGCAGCCCTCTGGCCGCCCTACATGAACCTGAGCTCAGC 910  
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 911 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCAATGCTGAACCTC 970  
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140  
Db 971 CTTGAGCCCACTCTGTGAGACAAGCTTCTAGTTAGAGACGCTTGTGATGAATGATGAG 1030  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
Db 1031 GAGGAACCTGTGCAATTTGAAGACAGAAACCGGATTTGCTGTCAGAAACAATGGAAT 1090  
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTyrPheSerAla 180  
Db 1091 GAATCAGGTGTAAAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGTTCTCTGCA 1150  
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200

Db 1151 TTTCTGAATGTTCTTCGTCAAAACAGAAACAATGAACCTTGCCAAAGATTAAACAGGCTCT 1210  
Qy 201 AspCysSerG]uSerAsnA]aglul]leg]uAsnLeuSerG]nVal]AspG]yProG]nVal] 220  
Db 1211 GATTGCTCAGAAAGCAATGCAGAGATTGAGAATTATCACAAGTTGATGGTCTCAAGTG 1270  
Qy 221 G]ug]uG]nLeuLeuSerThrThrVal]G]nProAsnLeuG]ulysG]luVal]TrpG]lyMet 240  
Db 1271 GAAGAGCAACTCTTTCAACCAACAGTTCAGCCAAATCTGAGAAGAGGCTCGGGGCATG 1330  
Qy 241 G]uAsnAsnSerSerG]uSerSerPheAlaAspSerSerVal]SerG]uSerAspThr 260  
Db 1331 GAGAAATACTCATCAGAAATCACTTTTGAGATTCTTTCTGTAAGTTTCAAGATCAGACACA 1390  
Qy 261 SerLeuAlaG]uG]ySerVal]SerCysLeuAspG]uSerLeuG]yHisAsnSerAsnMet 280  
Db 1391 AGTTTGGCAGAGAGAGAGTGTCACTGCTTAGATGAAGTCTTGACATAACAGCAACATG 1450  
Qy 281 G]ySerAspSerG]yThrMetC]ySerAspSerAspG]uG]uAsnVal]AlaAlaArgAla 300  
Db 1451 GGCAGTGATTCAGGCACCATGGGAAGTGATTCAGATGAAGAATGTGGCAGCAAGAACA 1510  
Qy 301 SerProG]uProG]uLeuG]nLeuArgProTyrg]lnMetG]luVal]AlaG]nProAlaLeu 320  
Db 1511 TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCCAAGCCTTG 1570  
Qy 321 G]ug]yLyAsnIleIleCysLeuProThrG]ySerG]yLyThrArgVal]AlaVal] 340  
Db 1571 GAAGGAGAAATATCATCATCTGCTCCCTACAGGAGTGGAAGAACAGAGTGGCTGTT 1630  
Qy 341 TyrIleAlaLyAspHisLeuAspR]yS]yS]yVal]AserG]uProG]yLyS]yVal]Ile 360  
Db 1631 TACATTGCCAAGGATCATTGACAAAGAAAAAGCATCTGAGCCTGAAAAAGTTATA 1690  
Qy 361 ValLeuVal]AsnLyS]yVal]LeuVal]G]uG]nLeuPheArg]yS]yG]uPheG]nProPhe 380  
Db 1691 GTTCTTGTCATAAGTACTGCTAGTTGAACAGCTCTCCGCAAGAGTTCCAAACCATTT 1750  
Qy 381 LeuLyS]yS]yTrpTyrg]nVal]IleG]yLeuSerG]yAspThrG]nLeuLyS]yIleSerPhe 400  
Db 1751 TTGAAGAAATGSTATCGTGTATTGGAATTAAGTGGTGATACCCAACTGAATAATCATTT 1810  
Qy 401 ProG]uVal]Val]ySerCysAspIleIleIleSerThrAlaG]nIleLeuG]uAsnSer 420  
Db 1811 CCAGAAAGTTGTCAAGTCCGTGTATATTATTCAGTACAGCTCAAACTCTGAAAACTCC 1870  
Qy 421 LeuLeuAsnLeuG]uAsnG]yG]uAspAlaG]yVal]G]nLeuSerAspPheSerLeuIle 440  
Db 1871 CTCTTAAACTTGAAAAATGGAGAAGATGCTGTGTTCAAATTGTCAAGACTTTTCCCTCATTT 1930  
Qy 441 IleIleAspG]uCysHisHisThrAsnLyS]yVal]Val]Tyrg]nAsnIleMetArgHis 460  
Db 1931 ATCATTTGATGAATGTCATCACCAACAAGAACAGAGTATTAATACATCATGAGGCAT 1990  
Qy 461 TyrLeuMetG]nLyS]yLeuLyAsnAsnArgLeuLyS]yG]uAsnLyS]yProVal]IlePro 480  
Db 1991 TATTTCATGCAGAGTTGAATAACAATAGACTCAAGAAAGAAAAACAACACAGTATCCC 2050  
Qy 481 LeuProG]nIleLeuG]yLeuThrAlaSerProG]yVal]G]yG]yAlaThrLyS]yAla 500  
Db 2051 CTTCTCTCAGATACTGGGACTAACAGCTTCACTGTGTGGAGGGGCCACGAGCAAGCC 2110  
Qy 501 LysAlaG]uG]uHisIleLeuLyS]yLeuCysAlaAsnLeuAspAlaPheThrIleLyS]yThr 520  
Db 2111 AAAGCTGAAGAACACATTTTAAACTATGTGCCAAATCTTGATGCATTACTATTAAAACT 2170  
Qy 521 ValLyS]yAsnLeuAspG]nLeuLyS]yAsnG]nIleG]nG]uProCysLyS]yS]yPheAla 540  
Db 2171 GTTAAAGAAAACTTGATCAACTGAAAAAACAATAACAGAGCCATGCAAGAAAGTTGCC 2230  
Qy 541 IleAlaAspAlaThrArgG]uAspProPheLyS]yLyS]yLeuG]uIleMetThrArg 560

Db 2231 ATTGCAGATGCACACAGAGAGATCCATTTAAGAGAACTTCTAGAAATAATGACAAAG 2290  
Qy 561 IleG]nThrTyrg]lnMetSerProMetSerAspPheG]yThrG]nProTyrg]uG]n 580  
Db 2291 ATTCAACTTATTGTCTCAATGAGTCCCAATGTCTGATTTTGGAACTCAACCTATGAACAA 2350  
Qy 581 TrpAlaIleG]lnMetG]uLyS]yS]yAlaAlaLyS]yG]yAsnArg]yS]yG]uArgVal]Cys 600  
Db 2351 TGGGCTATTCAATGGAAAAAAGCTGCMAAGAAAGAAATCGCAAGAAACGTGTTGT 2410  
Qy 601 AlaG]uHisLeuArg]yS]yTyrg]nG]uAlaLeuG]nIleAsnAspThrIleArgMetIle 620  
Db 2411 GCAGAACATTTGAGGAAGTACAAATGAGGCCCTACAAATTAAATGACACAAATTCGAATGATA 2470  
Qy 621 AspAlaTyrg]nHisLeuG]uThrPheTyrg]nAsnG]uLyS]yAspR]yS]yPheAlaVal] 640  
Db 2471 GATGCGTACTCATCTTGAAACTTTCTATTATGAAGAGAAAGATTAAGATTGCAGTC 2530  
Qy 641 IleG]uAspAspSerAspG]uG]yG]yAspAspG]uTyrg]nAspG]yAspG]uAspG]u 660  
Db 2531 ATAGAAAGATGATAGTAGAGGGTGTGATGATGAGTATTGTGATGGTATGAAGATGAG 2590  
Qy 661 AspAspLeuLyS]yS]yProLeuLyS]yLeuAspG]uThrAspArgPheLeuMetThrLeuPhe 680  
Db 2591 GATGATTTAAAGAAACCTTTGAAACTGGATGAACACAGATAGATTCTCATGACTTTATTT 2650  
Qy 681 PheG]uAsnAsnLyS]yMetLeuLyS]yArgLeuAlaG]uAsnProG]uTyrg]uAsnG]uLyS]y 700  
Db 2651 TTTGAAAACAATAAAATGTGAAAAAGCTGGCTGAAAAACCAAGAAATATGAATAAG 2710  
Qy 701 LeuThrLyS]yLeuArgAsnThrIleMetG]uG]nTyrg]nThrArgThrG]uG]uSerAlaArg 720  
Db 2711 CTGACCAAAATTAAGAAATACCATATGAGACAAATATACTAGACTGAGGAATCAGACAGA 2770  
Qy 721 G]yIleIlePheThrLyS]yThrArgG]nSerAlaTyrg]nAlaLeuSerG]nTrpIleThrG]u 740  
Db 2771 GGAATAATCTTTACAAAACAGACAGAGTGCATATGCCGTTTCCAGTGATTACTGAA 2830  
Qy 741 AsnG]uLyS]yPheAlaG]uVal]G]yVal]yS]yAlaHisIleLeuIleG]yAlaG]yHisSer 760  
Db 2831 AATGAAAAATTTGCTGAAGTAGAGAGTCAAGGCCCACTCATTTGAGAGCTGACACAGC 2890  
Qy 761 SerG]uPheLyS]yProMetThrG]nAsnG]uG]nLyS]yG]uVal]IleSerLyS]yPheArgThr 780  
Db 2891 AGTGAGTTCAAAACCATGACACAGAAATGAACAAGAAAGTCAATTAGTAAATTTCCGACT 2950  
Qy 781 G]yLyS]yIleAsnLeuIleAlaThrThrVal]AlaG]uG]uG]yLeuAspIleLyS]yG]u 800  
Db 2951 GGAAAAATAATCTGCTTATCGCTACACAGTGGCAGAAAGGCTGTGATATTAAAGAA 3010  
Qy 801 CysAsnIleVal]IleArgTyrg]yLeuVal]ThrAsnG]uIleAlaMetVal]G]nAlaArg 820  
Db 3011 TGTAACATTGTTATCCGTTATGGTCTGTCACCAATGAATAAGCCATGGTCCAGGCCGT 3070  
Qy 821 G]yArgAlaArgAlaAspG]uSerThrTyrg]nVal]LeuVal]AlaHisSerG]ySerG]yVal] 840  
Db 3071 GGTGAGCCAGAGCTGATGAGACACCTACGTCTGTGTTGCTCACAGTGGTTCAAGAGTT 3130  
Qy 841 IleG]uHisG]uThrVal]AsnAspPheArgG]uLyS]yMetMetTyrg]yS]yAlaIleHisCys 860  
Db 3131 ATGCAAATGAGACAGATTATGATTTCCGAGAGAAAGATGATGATTAAGCTATATCATTTGT 3190  
Qy 861 ValG]nAsnMetLyS]yProG]uG]uTyrg]nAlaHisLyS]yIleLeuG]uLeuG]lnMetG]nSer 880  
Db 3191 GTTCAAAATATGAACAACAGAGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAAAGT 3250  
Qy 881 IleMetG]uLyS]yMetLyS]yThrLyS]yArgAsnIleAlaLyS]yHisTyrg]yS]yAsnAsnPro 900  
Db 3251 ATAATGGAAGAAAGAAATGAACAACAGAGAAATATGCCCAAGCATTACAAAGATTAACCCA 3310  
Qy 901 SerLeuIleThrPheLeuCysLyS]yAsnCysSerVal]LeuAlaCysSerG]yG]uAspIle 920  
Db 3311 TCACATAATACTTTCTTTGCAAAAACTGCAGTGTGTAGCCTGTTCTGGGGAAGATATC 3370

Oy 921 HisValIleGluLysMethHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
|||  
Db 3371 CATGTAATGAGAAAATGCATCACTCAATATGACCCCAATTCAGAGAACTTTACATT 3430  
  
Oy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
|||  
Db 3431 GTAGAGAAAAACAAAACACTGCAGAAAGTGTGCCGACTATCAATAATGATGTAATC 3490  
  
Oy 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
|||  
Db 3491 ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGGCCT 3550  
  
Oy 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000  
|||  
Db 3551 TGTCTCAAAATAAGGAATTTGTAGTGGTTTCAAAATAATTCACAAAGAAACAATAC 3610  
  
Oy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGlyCysLeu 1020  
|||  
Db 3611 AAAAAGTGGTAGAATTACCTATCAATTTCCCAATCTTGACTATTCAGATGCTGTTTA 3670  
  
Oy 1021 PheSerAspGluAsp 1025  
|||  
Db 3671 TTATGATGAGGAT 3685  
  
RESULT 9  
ABA04908  
ID ABA04908 standard; cDNA; 3372 BP.  
XX  
AC ABA04908;  
XX  
DT 05-MAR-2002 (first entry)  
XX  
DE Human RNA helicase RH16 coding sequence.  
XX  
KW Human; RH16; RNA helicase; cytosolic; virucide; anti-HIV;  
KW immunosuppressive; immunostimulatory; antirheumatic; antiarthritic;  
KW antiarteriosclerotic; osteopathic; antidiabetic; hepatotropic;  
KW antiinflammatory; cancer; infection; HIV; hepatitis; genetic disease;  
KW autoimmune disease; graft rejection; vaccine; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key location/Qualifiers  
FT CDS 155..3232  
FT /\*tag= a  
FT /product= "Human RH16"  
XX  
PN WO200185955-A1.  
XX  
PD 15-NOV-2001.  
XX  
PF 11-MAY-2001; 2001WO-FR001441.  
XX  
PR 11-MAY-2000; 2000FR-0006030.  
XX  
PA (ISTA-) ISTAC.  
XX  
PI (INSP ) INST PASTEUR LILLE.  
XX  
PI Bahr G, Cocude C, Capron A;  
XX  
DR WPI: 2002-082898/11.  
DR P-PSDB; AAM47798.  
XX  
PT New polypeptide, useful for treating and diagnosing cancer or  
PT inflammation, and drug screening, comprises a human polynucleotide  
PT homologous to RNA helicase.  
XX  
PS Claim 7; Page 85-89; 114pp; French.  
XX  
CC The present sequence is the coding sequence for human RH16. RH16 is a  
CC 116kDa protein and has homology to RNA helicases (DEXH box). RH16 and  
CC its coding sequence are useful for treating cancer; acute or chronic

CC infections (especially by HIV or hepatitis B or C); inherited genetic  
CC diseases; (auto)immune diseases (particularly rheumatism, arthritis, and  
CC arteriosclerosis, osteoporosis and diabetes, but many others listed) and  
CC to prevent graft rejection. RH16 and its coding sequence are also useful  
CC for inducing, or increasing, the immune response to a vaccine  
XX  
SQ Sequence 3372 BP; 1157 A; 637 C; 748 G; 830 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 0 Length: 3372  
Score: 5285.00 Matches: 1020  
Percent Similarity: 99.7% Conservative: 2  
Best Local Similarity: 99.5% Mismatches: 3  
Query Match: 99.5% Indels: 0  
DB: 6 Gaps: 0  
  
US-09-515-363C-2 (1-1025) x ABA04908 (1-3372)  
  
Oy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
|||  
Db 155 ATGTGCAATGGGTATTCACAGACGAGAAATTCGCTATCTCATCTGCTTCAGGGCC 214  
  
Oy 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
|||  
Db 215 AGGGTGAATAATGTACATCCAGGTGAGCCCTGTGACTACCTGACCTTCTGCCTGCA 274  
  
Oy 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
|||  
Db 275 GAGGTGAAGAGACAGATTCAAGAGACAGTCCGCACTCCGGAACTGACAGGAGTTGAA 334  
  
Oy 61 LeuLeuLeuSerThrLeuGluLysGlyValTyrHisLeuGlyTyrThrArgGluPheVal 80  
|||  
Db 335 CTGCTGTGAGCACCTTGAGAGAAAGGAGTCTGGCACTTGTTGACTCGGGAATTCGTG 394  
  
Oy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
|||  
Db 395 GAGGCCCTCCGAGAACCGGACGCCCTGTGGCCGCCCTACATGAACCTGAGCTCAGC 454  
  
Oy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
|||  
Db 455 GACTTGCCCTCTCCATGCTTTGAGAACGCTGATGATATATCTCCAATGCTGAACCTC 514  
  
Oy 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140  
|||  
Db 515 CTTCAAGCCCACTCTGTGACCAAGCTTCTAGTAGAGAGCTTGTGATAGTGATGAGAG 574  
  
Oy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaAlaGluAsnAsnGlyAsn 160  
|||  
Db 575 GAGGAATCTGTGACATTTGAAGACAGAAACCGGATTGCTGTCGAGAAACAATGGAAT 634  
  
Oy 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180  
|||  
Db 635 GAATCAGGTGTAAGAGAGTAACTAAAAAGGATTGTGCAGAAAGAAAACCTGGTCTCTGCA 694  
  
Oy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
|||  
Db 695 TTTCTGAATGTTCTTCGTCAAACAGGAAACAACTGTCTCCAAAGAGTTAAACAGGCTCT 754  
  
Oy 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
|||  
Db 755 GATTGCTCAGAAAGCAATGAGAGATTGAGATTATACAAAGTTGATGCTCCCAAGTG 814  
  
Oy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240  
|||  
Db 815 GAAGAGCAACTCTTTTCAACCAAGTTCAAGCCAAATCTGGAGAGAGGCTGGGGCATG 874  
  
Oy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
|||  
Db 875 GAGATTAATCTCATCAGAAATCATCTTTGACAGATTCTTCTGTAGTTTCAGAAATCAGACACA 934  
  
Oy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
|||  
Db 935 AGTTTGGCAGAGAGAGAGTGTCAAGCTGTAGATGAAGTCTTGGACATTAACAGCAACATG 994

QY	281	GLYSerAspSerGlyThrMetGlySerAspSerAspGluAsnValAlaAlaArgAla	300
Db	995	GGCAGTGAATTCAGGCACCATGGGAAGTGAATTCAGATGAAGAGAATGTGGCAGCAGACGA	1054
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
Db	1055	TCCCCGAGCCAGCACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCACGCCCTTG	1114
QY	321	GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
Db	1115	GAAAGGAGAATATCATCATCTGCTCCCTACAGGAGCTGAAAAACAGAGTGGCTGT	1174
QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle	360
Db	1175	TACATTGCCAAGATCACTTAGACAAGAAAAACATCTGAGCCTGAAAAAGTTATA	1234
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380
Db	1235	GTTCTTGTCATTAAGTAAGTACTGCTAGTTGACAGCTCTTCGCAAGAGTTCCAACTTT	1294
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1295	TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGATACCAACTGAAATATCATTT	1354
QY	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1355	CCAGAAGTGTCAAGTCTGTGATATTATTATCAGTACAGCTCAATCCTTGAAACTCC	1414
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1415	CTCTTAACTTGGAAAAATGGAAGAGATGCTGGTGTTCAAATTGTCAGACTTTCTTCATT	1474
QY	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1475	ATCATTGATGAATGTCATCACACCAACAAGAGCAGTGTATTAATACATCATGAGCAT	1534
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
Db	1535	TATTGATGCAGAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACAGTGAATCCC	1594
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500
Db	1595	CTTCCTCAGATACTGGGACTAACAGCTTCACTGGTGTGGAGGGGCCACGAAGCAACC	1654
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1655	AAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACT	1714
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	1715	GTTAAAGAAAAACCTTGATCAACTGAAAAACCAATACAGAGCCATGCAAGAAGTTGCC	1774
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
Db	1775	ATTGCAGATGCACACGAGAGATCCATTTAAAGAGAACTTCTAGAATATGACACAGG	1834
QY	561	IleGlnThrTyrCysGlnMetSerPrometSerAspPheGlyThrGlnProTyrGluGln	580
Db	1835	ATTCAACTTATTGTCAATGAGTCCCAATGTCAGATTTGGAACTCAACCCCTATGACAA	1894
QY	581	TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1895	TGGGCCATTCAAAATGAAAAAAAGCTGCAAAAGAAAGAAATCGCAAAAGAAAGTTTGT	1954
QY	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	1955	GCAGAACATTTGAGGAAGTACATAAAGCCCTACAATTAATGACACAATTCGAATGATA	2014
QY	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal	640
Db	2015	GATCGTATACTCATCTTGAACTTTCTATTAATGAAGAGAAAGATTAAGAAGTTTGCACTC	2074

OY		641	IIleGIUAAspASpSerAaspGluGIyGLyAAspAsgIuTYrCYAspAsgIYAAspGluAAspGlu	660
Db		2075	ATAGAAGATGATGACTGATGAGGGTGGTGATCATGATTGTGATGGTGATGAAGATGAG	2134
OY		661	AAspAsgLeuLYsLYsProlEuLYsLeuAAspGluThrAAspArgPheLeuMetThrLeuPhe	680
Db		2135	GATGATTTAAAGAAACCTTTGAAAACGTGATGAACAAGATAGATTTCATGACTTTTAATT	2194
OY		681	PheGIUAsnAAsnLYsMetLeuLYsAsgLeuAglUAAsnProlGUtyrGIUAAsnGluLYs	700
Db		2195	TTTGAAAAACAATAAATCTTGAAAAGCGCTGGTGAALAACCAAGATATGAATAAGAAAAG	2254
OY		701	LeuThrLYsLeuAArgAAsnThriIeMetGluGlnTYrThrArgThrgIuGluSeraIAArg	720
Db		2255	CTGACCMAATTAAAGAAATACCATATGAGCAATATACTAGACTGAGGAATCAGCACGA	2314
OY		721	GIyIleilePheThrLYsThrrArgGlnSeraIATyrAlaleuSergIIntPriIeThrgIu	740
Db		2315	GGAATAATCTTACAAAAACACGACAGAGTGCAATATGCCCTTCCAGTGATTAAGTCAA	2374
OY		741	AsnGIULysPheAglagluValgLyvalLYsAlahisIsleuIlegIyAlagIyhissEr	760
Db		2375	AATGAAAAATTGCTGAAGTAGAGTCAAGCCCAACCATCTGATTGAGCTGCACACAGC	2434
OY		761	SerGIuPheLYsPrometThrgInAsnGIuGlnLYsGIuValIleserLYsPheArgThr	780
Db		2435	AGTGAGTTCAAAACCATGACACAGAAATGAACAAAAGAAGTCATTAGTAATTTCCACT	2494
OY		781	GIlyLYsIleAsnLeuLeuIlealathrThrValAlagIuGluGlyLeuAAspIleLYsGIu	800
Db		2495	GGAATAATAATCTGCTTATCGCTACCAAGTGCAGAGAGAGGTCTGGATATTAAAGAA	2554
OY		801	CysAAsnIleValIleargTYrGIyleuValThraAsnGIuIleAlametValGlnAlaArg	820
Db		2555	TGTAACATTGTATCCGTTATGCTCGTCAACCAATGAATAAGCCATGATGCCAGCCCCGT	2614
OY		821	GIyAArgAlaArgIAAspGluSerThrTYrValleuValAlahissErgLYsErgLYsVal	840
Db		2615	GGTCGAGCCAGAGCTGATGAGAGCACCTTACGTCCTGCTCACAAGTGGTTCAGAGATT	2674
OY		841	IleGIuHIsgIuThrValAsnAspPheArgGIuLYsMetMetTYrLYsAlaleHisCys	860
Db		2675	ATCGAACGTGAGACGTTAATGATTTCCGAGAGAAATGATGATATAAGCTATACATTGT	2734
OY		861	ValGlnAsnMetLYsProlGIuGluTYrAlahisLYsIleleuGIuleuGlnMetGlnSer	880
Db		2735	GTTCAAAATATGAACAAGAGAGATATGCTCATAGATTTTGGAATTACAGATGCAAGT	2794
OY		881	IleMetGIuLYsLYsMetLYsThrLYsAArgAAsnIleAlalyshISTYrLYsAAsnPro	900
Db		2795	ATAATGAAAAAGAAAAATGAAAAACCAAGAGAAATATTGCCAAGCATTACAAGATTAACCCA	2854
OY		901	SerLeuIlethrPheLeuCysLYsAAsnCysSerValleuAlaCyserGIyleuAAspIle	920
Db		2855	TCACTAATTAACCTTCTTTCGCAAAAACGTGAGTGTAGCCCTGTTCTGGGAGATATC	2914
OY		921	HISValIleGIuLYsMetHisHISValAsnMetThrpRolGIuPheLYsGIuLeuTYrIle	940
Db		2915	CATGTAATTGAGAAATGCATCAGCTCAATATGACCCCAGAAATTCAGAAGAACTTTACATT	2974
OY		941	ValAArgGIUAsnLYsAlaleuGlnLYsLYsCyseAlaAspTYrGlnIleAsnGIyleuIle	960
Db		2975	GTAAGAGAAAAACAAGCACTGCCAAAAGAGTGTGCCGACTATCAATAATATGTTGAATC	3034
OY		961	IleCYsLYsCYsGIyleGlnAlaTrPGlyThrmecMetValhisLYsGIyleuAAspLeuPro	980
Db		3035	ATCTGCAAAATGTGCCAGGCTTGGGCAACAATGATGTGCACAAAGGCTTAGATTGGCT	3094
OY		981	CysLeuLYsIleArgAsnPheValValValPheLYsAAsnAsnSerThrLYsLYsGlnTYr	1000
Db		3095	TGTTCTCAAAATAAGAAATTTTGTAGTGTGTTTTCAAAAAATAATTCAACAAAGAAACAATAC	3154
OY		1001	LYsLYsTrpValGIuLeuProIleThrpPheProAsnLeuAAspTYrSergIuCyseLYsLeu	1020

Db 3155 AAAAAGGGGTAGATATACCTATCACAATTCCCAATCTTGACTATTCAGAAATGCTGTTA 3214  
Qy 1021 pheSeraspGluasp 1025  
Db 3215 TTTAGTCATGAGCAT 3229  
RESULT 10  
ADC30823  
ID ADC30823 standard; cDNA; 3446 BP.  
XX  
AC ADC30823;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human novel cDNA sequence, SEQ ID NO:905.  
XX  
KW Human; diagnostic; drug screening; forensics; gene mapping;  
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
KW ulcers; osteoporosis; autoimmune disease; cancer;  
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;  
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;  
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;  
KW gene therapy; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029271-A2.  
XX  
PD 10-APR-2003.  
XX  
PF 24-SEP-2002; 2002WO-US030474.  
XX  
PR 24-SEP-2001; 2001US-0324631P.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T,  
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;  
PI Haley-Vicente D, Drmanac RT;  
XX  
DR WPI; 2003-371981/35.  
DR P-PSDB; ADC31794.  
XX  
PT New polynucleotide and polypeptide useful for diagnosing, preventing or  
PT treating conditions such as neurodegenerative diseases, anemias, platelet  
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
PT cancer.  
XX  
PS Claim 1; SEQ ID NO 905; 1185bp; English.  
XX  
CC The invention relates to 971 novel human cDNA sequences (ADC29919-  
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
CC invention also relates to nucleic acid sequences over 99% identical with  
CC the novel human cDNAs. The invention additionally encompasses expression  
CC vectors and host cells comprising a nucleic acid of the invention; the  
CC recombinant production of a polypeptide of the invention; an antibody  
CC against a polypeptide of the invention; a method of detecting  
CC polynucleotides or polypeptides of the invention; and methods of  
CC identifying a compound which binds to a polypeptide of the invention. The  
CC invention further discloses methods of preventing, treating or  
CC ameliorating a medical condition; kits comprising polynucleotide probes  
CC and/or monoclonal antibodies for carrying out the methods of the  
CC invention; methods for the identification of compounds that modulate the  
CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention  
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33394). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are

CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anaemia, platelet  
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a specifically  
CC claimed human cDNA sequence of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 3446 BP; 1184 A; 649 C; 772 G; 841 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 0 Length: 3446  
Score: 5285.00 Matches: 1020  
Percent Similarity: 99.7% Conservative: 2  
Best Local Similarity: 99.5% Mismatches: 3  
Query Match: 99.5% Indels: 0  
DB: 10 Gaps: 0  
  
US-09-515-363C-2 (1-1025) x ADC30823 (1-3446)  
Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
Db 237 ATGTGGAATGGGTATTCCACAGACGAAATTCCGCTATCTCATCTCGTCTCAGGGCC 296  
Qy 21 ArgValIysMetTyrIleGlnValGluProValIleuAspTyrLeuThrPheLeuProAla 40  
Db 297 AGGGTGAATAATGATCATCCAGGTGAGCGCTGTGCTGACTACTGACCTTTCTGCCTGCA 356  
Qy 41 GluValIysGluGlnIleGlnArgThrValAlaThrSergIysMetGlnAlaValGlu 60  
Db 357 GAGGTGAAGAGACAGATTCAAGACAGACAGTCCGCACTCCGGAAACATGACAGGCA GTTGA 416  
Qy 61 LeuLeuLeuSerThrLeuGluIysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80  
Db 417 CTGCTGCTGAGCACCTTGAGAAAGGAGTGTGGCACTTGGTGTGACTCGGAATTGCGTG 476  
Qy 81 GluAlaIeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
Db 477 GAGGCCCTCCGAGAACCGGACGCCCTTGCGCGCCGCTACATGAACCTGAGCTCAGC 536  
Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 537 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCAATGCTGAACCTC 596  
Qy 121 LeuGlnProThrLeuValAspIysLeuLeuValArgAspValIleuAspIysCysMetGlu 140  
Db 597 CTTCAGCCCACTCTGTGGACAAAGCTTCTAGTTAGAGACGCTTGTGATTAAGTCATGGAG 656  
Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
Db 657 GAGGAACCTGTTGACCAATTGAAGACAGAAACCGGATTGCTGTGCAGAAAACAATGGAAT 716  
Qy 161 GluSerGlyValArgGluLeuLeuIysArgIleValGlnIysGluAsnTrpPheSerAla 180  
Db 717 GAATCAGGTGTAGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAACGTGTTCTCTGCA 776  
Qy 181 PheLeuAsnValIleuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 777 TTTCTGAATGTTCTTCGTCAAAACAGAAACATGAACCTGTCCAAGAGTTAACAGGCTCT 836  
Qy 201 AspCysSergIuSerAsnAlaGluIleGluAsnLeuSergIuValAspGlyProGlnVal 220  
Db 837 GATTGCTCAGAAAGCAATGCAAGATTGAGAAATTATCACAAAGTTGATGTCCTCAAGTG 896  
Qy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIysGluValTrpGlyMet 240  
Db 897 GAAGAGCAACTTCTTTCAACACAGATTGAGCAAAATCTGAGAAAGAGGCTCTGGGGCATG 956  
Qy 241 GluAsnAsnSerSergIuSerSerPheAlaAspSerSerValValSergIuSerAspThr 260

Db 957 GAGATAACTCATCAGAACTCTTTTGAGATTCTTGTAGTTTCAGAAATCAGACACA 1016  
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 1017 AGTTGGCAGAGAAGAGTGTCACTGCTGTAGATGAAGCTTTGACATTAACAGCAACATG 1076  
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluAsnValAlaAlaArgAla 300  
Db 1077 GGCACGTGATTCAGGCACCACTGGGAAGTGATTCAAGATGAAGAATGTGGCAGCAAGACA 1136  
Qy 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
Db 1137 TCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCCTTG 1196  
Qy 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340  
Db 1197 GAAGGGAAGAATATCATCTATCTGCCCTCCCTACAGGGAGTGGAACCAACAGAGTGCTGTT 1256  
Qy 341 TyrIleAlaLysAspHisLeuAspLysLysLysValAserGluProGlyLysValIle 360  
Db 1257 TACATTGCCAAGGATCATTAGACAGAAGAAAAAGCATCTGAGCCTGGAAGTTATA 1316  
Qy 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380  
Db 1317 GTTCTTGTCATAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACCATTT 1376  
Qy 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400  
Db 1377 TTGAAGAAATGGTATCGTTATTGGAATTAAAGTGTGATACCCTCAACTGAAAAATATCATTT 1436  
Qy 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
Db 1437 CCAGAGTTGTCAAGTCCGTGATATTATTATCAGTACAGCTCAAAATCTTGAAAACTCC 1496  
Qy 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
Db 1497 CTCTTAAACTTGAAAAATGAGAGATGCTGTGTTCATTTGACACTTTTCCTTCATT 1556  
Qy 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460  
Db 1557 ATCATTTGATGATGTATCATCACCAACAAGAAGCAGTATATAATACATCATGAGGCAT 1616  
Qy 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480  
Db 1617 TATTTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCACTGATTCCC 1676  
Qy 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500  
Db 1677 CTTCCTCAGATACTGGGACTTAACAGCTTCACTGGTGTGGAGGGCCACGAGCAAGCC 1736  
Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
Db 1737 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACT 1796  
Qy 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540  
Db 1797 GTTAAAGAAAACTTGATCACTGAAAAAACCAATACAGAGCCATGCAAGAGTTTGCC 1856  
Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560  
Db 1857 ATTGCAGATGCAACCAAGAGAGATTCATTTAAAGAGAACTTCTAGAAATATGACAAAG 1916  
Qy 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580  
Db 1917 ATTCAAACTTATTGTCAAAATGAGTCCAATGTCAGATTTTGGAACTCAACCTTATGAACAA 1976  
Qy 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
Db 1977 TGGGCCATTCAAAATGAAAAAAAGCTGCAAAAAGAGGAATCGCAAAAGAAAGTGTGTGT 2036  
Qy 601 AlaGluHisLeuArgLysTyrArgGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620

Db 2037 GCAGAACATTTGAGGAAGTACATAAAGGCCCTACAAATTAATGACACAATTCGAATGATA 2096  
Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640  
Db 2097 GATCGGTATACTCATCTTGGAAACTTCTATATAGAAGAAAGATAGAAGTTTGCACTC 2156  
Qy 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 2157 ATAGAAGATGATAGTAGTAGCGGTGTGATGATGAGTATTGTGTGATGATGAAGATGAG 2216  
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2217 GATGATTTAAAGAACTTTGAAACTGGATGAAACAGATAGATTTCATGACTTATATT 2276  
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2277 TTGAAAAACAATAAATGTGAAAAAGCTGGCTGAAAAACCAAGATATGAAAAATGAAAG 2336  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720  
Db 2337 CTGACCAAAATTAAGAAATACCATAATGAGCAATATACTAGACTGAGGAATCAGACAGA 2396  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740  
Db 2397 GGAATTAATCTTACAAAAACACGACAGAGTGCATATGCCGCTTCCCACTGCAATTAAGAA 2456  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760  
Db 2457 AATGAAAAATTTGCTGAAGTAGAGAGTCMAAGCCACCATCTGATGTGAGCTGGACACAGC 2516  
Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2517 AGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAAGATCATTAATAATTTCGCACAT 2576  
Qy 781 GlyLysIleAsnLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2577 GGAATAATAATCTGCTTATCGCTACACAGTGGCAGAAAGAGTGTGATATTAAAGAA 2636  
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2637 TGTAACATTTGTTATCCGTTATGCTCTCGTCAACCAATGAATAGCCATGGTCCAGGCCGT 2696  
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2697 GGTGGAAGCCAGAGCTGATGAGAGCACCTACGTCCTGTTGCTCACAGTGTTCAGAGATT 2756  
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860  
Db 2757 ATCGAACGTGAGACAGTTAATGATTTCCGAGAGAAGATGATGATTAAGCTATACATTGT 2816  
Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2817 GTTCAAAATATGAACCGAGAGAGTATGCTCATTAAGATTTTGAATTAAGATGCAAGT 2876  
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900  
Db 2877 ATTAATGAAAAAGAAATGAATAAACCAAGAGAAATATTGCCAAGCATTAAGAATAACCCA 2936  
Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920  
Db 2937 TCACTTAATACTTTCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATC 2996  
Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
Db 2997 CATGTAATTGAGAAATGATCAAGTCAATATGACCCAGAAATTAAGGAACCTTACATT 3056  
Qy 941 ValArgGluAsnLysValLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 3057 GTAAGAGAAAAACAAGCACTGCAAAAAGAGTGTGCCGCACTATCAAAATTAATGTGAAATC 3116  
Qy 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3117 ATCTGCAATATGTGGCCAGGCTTGGGGAACAATGATGCTGCACAAAGGCTTAAGATTGCGCT 3176

QY 981 CysleuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000  
Db 3177 TGTCTCAAAATAAGGAATTTGTAGTGTTCAAAAATAATTCACAAAGAAACAATAC 3236  
QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLysLeu 1020  
Db 3237 AAAAAGTGGGTAGAAATTACCTATACATTCCCAATCTTGACTATTGAGAAATGCTGTTTA 3296  
QY 1021 PheSerAspGluAsp 1025  
Db 3297 TTTAGTGAATGAGGAT 3311

RESULT 11

ADN04879  
ID ADN04879 standard; cDNA; 3379 BP.  
XX  
AC ADN04879;  
DT 01-JUL-2004 (first entry)  
XX  
DE Antipsoriatic cDNA sequence #654.  
DE  
XX  
KM ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.  
XX  
OS Homo sapiens.  
XX  
PN WO2004028479-A2.  
XX  
PD 08-APR-2004.  
XX  
PF 25-SEP-2003; 2003WO-US030907.  
XX  
PR 25-SEP-2002; 2002US-0414006P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;  
PI Wu TD;  
XX  
DR WPI; 2004-305105/28.  
DR P-PSDB; ADN04880.  
XX  
XX  
PT New PRO nucleic acid or polypeptide, useful for preparing a  
PT pharmaceutical composition for diagnosing or treating psoriasis in a  
PT mammal.  
XX  
XX  
PS Claim 1; SEQ ID NO 1273; 3069pp; English.  
XX  
XX  
CC The invention relates to novel polynucleotide and polypeptides for  
CC treating psoriasis or a sequence having at least 80% identity to the  
CC above sequences. The nucleic acid is useful for preparing a composition  
CC for diagnosing or treating psoriasis in a mammal. This sequence  
CC corresponds to one of the polynucleotides of the invention.  
XX  
SQ Sequence 3379 BP; 1152 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0 Length: 3379  
Score: 5283.00 Matches: 1024  
Percent Similarity: 99.9% Conservative: 0  
Best Local Similarity: 99.9% Mismatches: 1  
Query Match: 99.5% Indels: 1  
DB: 12 Gaps: 0

US-09-515-363C-2 (1-1025) x ADN04879 (1-3379)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
Db 169 ATGTGCAATGGGTATTCACAGAGAAATTCGCTATCTCATCTCGTCTCAAGGCC 228  
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40

Db 229 AGGGTGAATAATGTACATCCAGGTGAGCCCTGTGCTGACTACCTGACCTTTCTGCCCTGCA 288  
QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 289 GAGGTGAAGAGACAGATTCAAGAGACAGTCCGCACTCCGGAAACATGACGAGTTGAA 348  
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTyrTrpThrArgGluPheVal 80  
Db 349 CTGCTGCTGAGCACCTTGAGAGAGGAGTCTGGCACCTTGTTGACTCGGGAATTCTGTG 408  
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
Db 409 GAGGCCCTCCGAGAACCCGACGCCCTCTGGCCGCCCTACATGAACCTGAGCTCACG 468  
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGTAACCTC 528  
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140  
Db 529 CTTACGCCCACTCTGGTGACAAAGCTTCTAGTTAGAGACGCTTGATTAAGTGCATGAG 588  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
Db 589 GAGGAACCTGTTGACAAATTGAAGACAGAAACCGGATTGCTGTCAGAAAACAATGGAAT 648  
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180  
Db 649 GAATCAGGTGTAAGAGAGCTACTAAAAAGATTGTGCAGAAAGAAACTGTTCTCTGCA 708  
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 709 TTTCTGAATGTTCTTCGTCAAACAGAAACAATGAACCTGTCCAAAGTTTAACAGGCTCT 768  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 769 GATTGCTCAGAAAGCAATGACAGAGATTGAGAAATTATCACAAAGTTGATGCTCTCAAGTG 828  
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240  
Db 829 GAAGAGCAACTCTTTTCAACCAAGTTCAGCCAAATCTGGAGAGAGGCTGGGGCATG 888  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValSerGluSerAspThr 260  
Db 889 GAGATTAATCATCATCATCTTTTGCAGATTCTTGTAGTTTCAGAAATCAGACACA 948  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 949 AGTTTGGCAGAAAGAGTGTCAAGCTGCTTAGATGAAAGTCTTGAGACATRAACGCAACATG 1008  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
Db 1009 GGCAGTGATTCAGGCACCATGGGAAGTGAATGAAAGAAATGTGGCAGCAAGAGCA 1068  
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
Db 1069 TCCCCGGAGCCAGAACTCAGCTCAGGCCCTTACCAATGGAAGTTGCCAGCCAGCCTTG 1128  
QY 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340  
Db 1129 GAAGGGAAGAATATCATCTGCTCCCTACAGGGAGTGAAAAACAGAGTGCTGTT 1188  
QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360  
Db 1189 TACATTGCCAAGGATCACTTAGACACAGAGAAAAAGCATCTGAGCCCTGAAAAAGTTATA 1248  
QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProphe 380  
Db 1249 GTTCTTGTCAATAAGTACTGCTAGTTGAACAGCTCTTCGCAAGGAGTTCCAAACCATTT 1308  
QY 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400  
Db 1309 TTGAAGAAATGTATCGTGTATTGATTAAGTGTGATACCCCAACTGAAAAATATCATTT 1368

QY	401	ProGluValValIySerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1369	CCAGAGTTGTCAGAGCTCTGTGATATTATTATCAGTACAGCTCAAACTCTTGAAAACCTCC	1428
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1429	CTCTTAACCTTGAAATGGAAGAATGCTGTGTTCATATGTCAGACTTTTCCCTCATTT	1488
QY	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1489	ATCATTTGATGATGTCTATCACACCAACAA - GAAGCAGTGTATAATACATCATGAGCAT	1547
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
Db	1548	TATTGTATGACAGAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACACAGTATTTCC	1607
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500
Db	1608	CTTCCTCAGATACCTGGGACTAACAGCTTCACCTGTGTGTGAGGGGCCACGAAGCAACC	1667
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1668	AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACT	1727
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	1728	GTTAAAGAAAACTTGATCAACTGAAAAACCAATACAGAGCCATGCAAGAGTTTGCC	1787
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGlnIleMetThrArg	560
Db	1788	ATTGACAGATGCMAACGAGAGAGATCCATTTAAAGAGAACTTCTAGAAATTAATGACAAAG	1847
QY	561	IleGlnThrTyrCysGlnMetSerPrometSerAspPheGlyThrGlnProTyrGlnGln	580
Db	1848	ATTCAACTTATTGTCAATGAGTCCCAATGTCAGATTTTGAAACTCAACCTATGAAACA	1907
QY	581	TrpAlaIleGlnMetGlnLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1908	TGGGCCATTCAATGGAAAAAAGCTGCAAAAAAGAGAAATCGCAAGAACGTGTTGT	1967
QY	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	1968	GCAGAACATTTGAGGAGTACCAATGAGGCCCTACAAATTAATGACACAAATTCGATGATA	2027
QY	621	AspAlaTyrThrHisLeuGlnThrPheTyrAsnGluGluLysAspLysPheAlaVal	640
Db	2028	GATGCGTACTCATCTTGAACTTCTATATGAGAGAAAGATAAGAGTTTGCACTC	2087
QY	641	IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	2088	ATAGAAAGATGATAGTGAAGGCTGCTGATGATGAGTATTGTGATGGTGAAGATGAG	2147
QY	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	2148	GATGATTTTAAAGAACTTTGAACTGGATGAACAAGATGATTTCTCATGACTTATTT	2207
QY	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db	2208	TTTGAAAACAATAAATGTTGAAAAGGCTGGCTGAAAAACCAAGATATGAAAATGAAAAAG	2267
QY	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg	720
Db	2268	CTGACCAAAATTAAGAAATACCAATATGAGCAATATACTAGGACTGAGGAATCAGACGA	2327
QY	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu	740
Db	2328	GGAATAATCTTTACAAAAACGACAGAGTGCAATATGCGCTTCCACAGTGAATTACTGA	2387
QY	741	AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer	760
Db	2388	AATGAAAAATTTGCTGAAGTAGGAGTCAAAAGCCCAACCATCTGATTTGAGCTGGACACAGC	2447

Qy	761	SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr	780
Db	2448	AGTGAATTCAAAACCATGACACAGAAATGAACAAAAGAGTCAATTAGTAAATTTCCGACT	2507
Qy	781	GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu	800
Db	2508	GGAAAAATCAATCTGCTTATCGCTACACAGTGGCAGAAGAGGTCTGGATATTAAAGAA	2567
Qy	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	820
Db	2568	TGTACATGTGTATCCGTTATGCTCTCGTCAACAATGAATAGCCATGTGTCAGGCCGT	2627
Qy	821	GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal	840
Db	2628	GCTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGCTGCTCACAGTGGTTCAGAGATT	2687
Qy	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860
Db	2688	ATCGAACATGACAGACGTTAATGATTTCCGACAGAAAGATGATGATTAAGCTATACATTGT	2747
Qy	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer	880
Db	2748	GTTCAAAATATGAACACAGAGAGTATGCTCATTAAGATTTTGGAAATTACAGATGCAAAAGT	2807
Qy	881	IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro	900
Db	2808	ATAATGAAAAGAAAATGAAAACCAAGAGAAATATTGCCAAGCATTACAAAGAATAACCCA	2867
Qy	901	SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle	920
Db	2868	TCACATAATACTTCTTTCGAAAACTGCAGTGCTAGCCTGTTCTGGGGAAGATATC	2927
Qy	921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheGluLeuTyrIle	940
Db	2928	CATGTAATTGAGAAAATGCATCACGTCAATATGACCCCAGAATTCAAGGAACTTTACATT	2987
Qy	941	ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle	960
Db	2988	GTAAGAGAAAACAAACCACTGCACAAAAGAGTGTCCGACTATCAAAATAAATGGTGAAATC	3047
Qy	961	IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro	980
Db	3048	ATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAAGGCTTAGATTGCGCT	3107
Qy	981	CysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnTyr	1000
Db	3108	TGTCTCAAAATAAGGAATTTTGTAGTGTTTTCAAAAATAATTCAACAAAGAAACAATAC	3167
Qy	1001	LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLysLeu	1020
Db	3168	AAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAGAATGCTGTTTA	3227
Qy	1021	PheSerAspGluAsp	1025
Db	3228	TTTAGTGATGAGGAT	3242
RESULT 12			
AAD11170 standard; DNA; 3131 BP.			
XX	AAD11170;		
XX	AC		
XX	DT 29-NOV-2001 (first entry)		
DE	Human melanoma differentiation associated-5 protein-related DNA.		
XX	Human melanoma differentiation associated gene; Mda-5; interferon; IFN;		
KW	RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;		
KW	neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;		
KW	breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;		
KW	central nervous system; cytosstatic; apoptosis; ds.		
OS	Homo sapiens.		

XX WO200164707-A1.  
PN  
XX 07-SEP-2001.  
PD  
XX 28-FEB-2001; 2001WO-US006960.  
PF  
XX 29-FEB-2000; 2000US-00515363.  
PR  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
PA  
XX Fisher PB, Kang D, Gopalkrishnan RV;  
PI  
XX WPI; 2001-565494/63.  
DR  
XX  
PT Nucleic acid sequences encoding a Melanoma Differentiation Associated  
PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral  
PT activity.  
XX  
XX Disclosure; Page 134-148; 152pp; English.  
PS  
XX  
XX The present invention relates to an isolated nucleic acid encoding a  
CC melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5  
CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.  
CC Mda-5 is a novel interferon (IFN) inducible gene with structural  
CC similarities to RNA helicases and CARD motif containing proteins. Mda-5  
CC is induced during terminal differentiation in human melanoma cells  
CC treated with the combination of recombinant fibroblast IFN and the  
CC antileukaemic compound mezerein (MEZ). Mda-5 is useful for identifying  
CC compounds that may induce its expression. Mda-5 is useful for treating  
CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma  
CC multiforme, cervical cancer, breast cancer, colon cancer, prostate  
CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a  
CC cancer of the central nervous system and apoptosis. The Mda-5 promoter  
CC exhibits melanocyte tissue specificity and minimises systemic toxicity.  
CC The present sequence is designated SEQ ID NO:2 in the sequence listing.  
CC present sequence is designated SEQ ID NO:2 in the sequence listing, but  
CC does not correspond to the sequence designated SEQ ID NO:2 in the main  
CC body of the specification (AAB10155). The present sequence is not further  
CC referred to in the specification, and has been represented in a protein  
CC format in the sequence listing  
XX  
SQ Sequence 3131 BP; 1075 A; 583 C; 699 G; 774 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 0 Length: 3131  
Score: 5276.50 Matches: 1024  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 0  
Query Match: 99.4% Indels: 3  
DB: 4 Gaps: 1  
  
US-09-515-363C-2 (1-1025) x AAD11170 (1-3131)  
QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
DB 1 ATGTGCAATGGGTATTCCACAGACGAGATTCCGCTATCTCATCTCGTCTCAGGGCC 60  
QY 21 ArgValIysMetTyrIleGlnValGluProValIleuAspTyrLeuThrPheLeuProAla 40  
DB 61 AGGTGAAATATGTACATCCAGGTGAGCCGTGCTGGACTAAGCTGACCTTTCTGCTGCA 120  
QY 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
DB 121 GAGGTGAAGGACGATTCAGAGGACAGTCCACCTCCGGGAAACATGCAGGCAATTGAA 180  
QY 61 LeuLeuLeuSerThrLeuGluIysGlyValTPrHisLeuGlyTPrThrArgGluPheVal 80  
DB 181 CTGCTGTGAGCACCTTGAGAGAGGAGTCTGGCACCTTGCTGGAATCGGAAATTCGTG 240  
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
DB 241 GAGGCCCTCCGAGAACCGGACGCCCTCTGCGCCGCCCTAATGAACCTGAGCTCACG 300

QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
DB 301 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACCTC 360  
QY 121 LeuGlnProThrLeuValAspIysLeuLeuValArgAspValLeuAspIysCysMetGlu 140  
DB 361 CTTGAGCCCACTCTGGTGACAAAGCTTCTAGTTAGAGACGCTTGATAGTGCATGCGAG 420  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
DB 421 GAGGAAGTGTGACAAATGGAAGACAGAAACCGGATTGCTGCGAGAAAACATGGAAT 480  
QY 161 GluSerGlyValArgGluLeuLeuIysArgIleValGlnIysGluAsnTrpPheSerAla 180  
DB 481 GAATCAGGTGTAAAGAGAGCTACTAAAAAGGATTGCGAGAAAAGAACTGGTCTCTGCA 540  
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
DB 541 TTTCTGAATGTTCTTCGTCAACAGGAACAATGAACTGTCCAGAGTTAACAGGCTCT 600  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAlaAspGlyProGlnVal 220  
DB 601 GATTGCTCAGAAAGCAATGCAGAGATTGAGAATTATACAAAGTTGATGGTCTCAAGTG 660  
QY 221 GluGlnGlnLeuLeuSerThrThrValGlnProAsnLeuGluIysGluValTPrGlyMet 240  
DB 661 GAAGAGCAACTCTTTCAACCAAGTTCCAGCCAAATCTGAGAAAGAGGTCTGGGGCATG 720  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValAlaSerGluSerAspThr 260  
DB 721 GAGATAACTCATCAGAAATATCTTTGCAGATTCTCTGTAGTTTCAGAAATCAGACACA 780  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
DB 781 AGTTGGCAGAGAGAAAGTGTACGCTGTTAGATGAAGTCTTGACATACGCAACATG 840  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
DB 841 GGCAGTATTCAGGCAACCATGGGAAGTATTGATGAAGAAGATGTGGCAGCAAGACA 900  
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
DB 901 TCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCAGCAGCCTTG 960  
QY 321 GluGlyIysAsnIleIleIleCysLeuProThrGlySerGlyIysThrArgValAlaVal 340  
DB 961 GAAGGAGAAATATCATCATCTGCTCCCTCCTACAGGAGTGGAAGAAACAGAGTGGCTGTT 1020  
QY 341 TyrIleAlaIysAspHisLeuAspIysIysIysAlaSerGluProGlyIysValIle 360  
DB 1021 TACATTGCCAAGATCATCTTAGACAGAGAAAAAGCATCTGAGCCTGAGAAAAATTATA 1080  
QY 361 ValLeuValAsnIysValLeuLeuValGluGlnLeuPheArgIysGluPheGlnProphe 380  
DB 1081 GTTCTTGTCAATAAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGAGTTCCAACCATTT 1140  
QY 381 LeuIysIysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuIysIleSerPhe 400  
DB 1141 TTGAAGAAATGTTATCGTTATTGATTAAAGTGTGATACCCTGAATAATATCATATT 1200  
QY 401 ProGluValValIysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
DB 1201 CCAGAAAGTGTCAAGTCCGTGAT--ATTATCAGTACAGCTCAAAATCCTTGAAAACTCC 1257  
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
DB 1258 CTCTTAAACTTGAAAATGGAAGAAGATGCTGTGTCAATTGTCAAGACTTTCCCTCATTT 1317  
QY 441 IleIleAspGluCysHisHisThrAsnIysGluAlaValTyrAsnAsnIleMetArgHis 460  
DB 1318 ATCATTTGATGATGTATCATCACCAACAAGAGCAAGTATATAATACATCATGAGCAT 1377

QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480  
|||  
Db 1378 TATTGATGCAGAGTTGAAAAACAATAGACTCAAGAAAAGAAAACCAACCACTGATTCCC 1437  
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAla 500  
|||  
Db 1438 CTTCCTCAGATACTGGGACTTAAACAGCTTCACCTGGTGTGGAGGGGCCACGAAAGCAAGCC 1497  
QY 501 LysAlaGlnGlnHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
|||  
Db 1498 AAAGCTGAAGAACACATTTTAAACTATGTGCCAACTCTGATGCATTACTATTAAAACT 1557  
QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540  
|||  
Db 1558 GTTAAAGAAAACTTGATCACTGAAAAACCAATACAGAGCCATGCAGAAAGTTTGC 1617  
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGlyLysLeuLeuGluIleMetThrArg 560  
|||  
Db 1618 ATTGCAGATGCACACAGAGAGATCCATTTAAAGAGAACTTCTAGAAATATGACAAAG 1677  
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlnGln 580  
|||  
Db 1678 ATTCAACTTATTTGTCAAATGAGTCCAATGTCAGATTGTGAACTCAACCTTATGAACAA 1737  
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
|||  
Db 1738 TGGGCCATTCAAAATGAAAAAAAGCTGCAAAAAAAGAAATCGCAAGAACGTGTGTGT 1797  
QY 601 AlaGlnHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
|||  
Db 1798 GCAGAACATTTGAGGAAGTACATGAGCCCTCAAAATTAATGACACAAATTCGAATGATA 1857  
QY 621 Asp-AlaTyrThrHisLeuGluThrPheTyrAsnGlnGluLysAspLysPheAlaVal 640  
|||  
Db 1858 GATCCGCGTATACTCATCTTGAAACTTTCTATATGAGAGAAAGATAAGAGTTTGACAG 1917  
QY 640 allIleGluAspAspSerAspGluGlyLysAspAspGluTyrCysAspGlyAspGluAspG 660  
|||  
Db 1918 TCATAGAGATGATAGTATGATGAGGGTGTGATGATGATGATGATGATGATGATGATG 1977  
QY 660 LysAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuP 680  
|||  
Db 1978 AGGATGATTTAAAGAAACCTTTGAAACTGGATGAACAGATAGATTTCTCATGACTTTAT 2037  
QY 680 hepheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluL 700  
|||  
Db 2038 TTTTGAACAAATAAATAGTTGTAAGCTGCTGAGAAACCCAGAAATATGAATAATGAAA 2097  
QY 700 yLeuThrLysLeuArgAsnThrIleMetGlnGlnTyrThrArgThrGluGlnSerAlaA 720  
|||  
Db 2098 AGCTGACCAATTAAGAAATACCAATATGAGCAATATTAAGACTGAGAAATCAGACAC 2157  
QY 720 rgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrG 740  
|||  
Db 2158 GAGGAATATCTTTACAAAAACAGCAGAGTGCAATATCGCCTTCCCACTGATTAAGT 2217  
QY 740 LysAsnGluLysPheAlaGluValGlyValLysAlaHisHisIleLeuIleGlyAlaGlyHis 760  
|||  
Db 2218 AAAATGAAAAATTGCTGAGTAGAGTCAAAAGCCACCATCTGATTGGAGCTGAGACACA 2277  
QY 760 erSerGluPheLysProMetThrGlnAsnGlnGlnLysGluValIleSerLysPheArgT 780  
|||  
Db 2278 GCAGTAGTTCAAAACCATGACACAGAAATGAACAAAAAGAAAGTCATTAGTAATTTGCA 2337  
QY 780 hrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGlnGlnGlyLeuAspIleLysG 800  
|||  
Db 2338 CTGAAAAATCAATCTGCTTATCGCTAACACAGTGGCAGAAAGAGTCTGATATTTAAAG 2397  
QY 800 lncYsaenIleValIleArgTyrGlyLeuValThrAsnGlnIleAlaMetValGlnAlaA 820  
|||  
Db 2398 AATGTAACATGTTTATCCGTTATGCTCTGCTCACCAATGAATAAGCCATGCTCGAGGCC 2457  
QY 820 rgGlyArgAlaArgAlaAspGlnSerThrTyrValLeuValAlaHisSerGlySerGlyV 840

Db 2458 GTGCTGAGCCAGAGCTGATGAGAGACACTTACGTCTGTTGCTCACAGTGGTTCAGAG 2517  
QY 840 allIleGlnHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisC 860  
|||  
Db 2518 TTATCGAAACATGAGACAGTAAATGATTTCCGAGAGAAAGATGATGATTAAGCTATACAT 2577  
QY 860 yValGlnAsnMetLysProGlnGluTyrAlaHisLysIleLeuGlnLeuGlnMetGlnS 880  
|||  
Db 2578 GTGTCAAAATATGAACCAAGAGAGTATGCTCATTAAGATTTTGGAAATTAACAGATGCAA 2637  
QY 880 erIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnP 900  
|||  
Db 2638 GTATTAATGAAAAAGAAATGAACCAAGAGAAATATTGCCAAGCATTAACAAGATAACC 2697  
QY 900 roSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyLysAspI 920  
|||  
Db 2698 CATCACTAATACTTCTTTGCAAAAACCTGCAGTGTCTAGCCTGTTCTGGGGAAGATA 2757  
QY 920 leHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrI 940  
|||  
Db 2758 TCCATGTAAATTGAGAAATGCATCACGTCAATATGACCCAGAAATTCAGAACTTTACA 2817  
QY 940 leValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluI 960  
|||  
Db 2818 TTGTAAGAGAAAAACAAGCACTGCACAAAAGAGTGTGCCGACTATCAATAAATGTTGAAA 2877  
QY 960 leIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuP 980  
|||  
Db 2878 TCATCTGCAAAATGTGGCCAGCTTGGGGAACATGATGTTGCACAAAGCCTTAGATTGC 2937  
QY 980 roCysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnT 1000  
|||  
Db 2938 CTGTCTCAAAATTAAGGAATTTTGTAGTGTGTTTCAAAAATATATCAACAAGAAACAT 2997  
QY 1000 yTrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysL 1020  
|||  
Db 2998 ACAAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTT 3057  
QY 1020 eupheserAspGluAsp 1025  
|||  
Db 3058 TATTTAGTGATGAGGAT 3074  
RESULT 13  
ADQ22237  
ID ADQ22237 standard; DNA; 3668 BP.  
XX  
AC ADQ22237;  
XX  
DT 26-AUG-2004 (first entry)  
XX  
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5057.  
KW soft tissue sarcoma; cytosstatic; gene therapy; vaccine; screening; human;  
KW ds.  
XX  
OS Homo sapiens.  
XX  
PN WO2004048938-A2.  
XX  
PD 10-JUN-2004.  
XX  
PF 26-NOV-2003; 2003WO-US038193.  
XX  
PR 26-NOV-2002; 2002US-0429739P.  
XX  
PA (PROT-) PROTEIN DESIGN LABS INC.  
XX  
PI Aziz N, Ginsburg WM, Zlotnik A;  
XX  
DR WPI; 2004-441208/41.  
XX  
PT Early detection of soft tissue sarcoma comprises determining expression

PT of a gene in a first soft tissue sample and a normal soft tissue sample  
PT and comparing the gene expression, also useful in treating soft tissue  
PT sarcoma.

XX Example 2; SEQ ID NO 5057; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma  
CC which comprises obtaining a first soft tissue sample from an individual  
CC and a normal soft tissue sample from the same or different individual,  
CC determining the expression of a gene in both samples and comparing the  
CC expression of the gene in both soft tissue samples, where a higher level  
CC of protein expression in the first soft tissue sample indicates the  
CC presence of soft tissue sarcoma. The method of the invention has  
CC cytosolic applications and may be useful for detecting soft tissue  
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic  
CC acid sequences may be useful in diagnostic and screening applications.  
CC The current sequence is that of a human soft tissue sarcoma-upregulated  
CC DNA of the invention. The current sequence is not shown within the  
CC specification per se but was submitted in CD format by the inventor.

XX Sequence 3668 BP; 1215 A; 730 C; 811 G; 872 T; 0 U; 40 Other;

Alignment Scores:

Pred. No.: 0 Length: 3668  
Score: 5251.00 Matches: 1015  
Percent Similarity: 99.1% Conservative: 1  
Best Local Similarity: 99.0% Mismatches: 9  
Query Match: 98.9% Indels: 0  
DB: 12 Gaps: 0  
US-09-515-363C-2 (1-1025) x ADQ22237 (1-3668)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
DB 169 ATGTGCAATGGGTATCCACAGACGAAATTCCGCTATCTCATCTCGTCTTCAGGGCC 228  
QY 21 ArgValIysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
DB 229 AGGGTGAATAATGTACATCCAGGTGAGCCTGTGCTGACTACCTGACCTTCTGCTGCA 288  
QY 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
DB 289 GAGGTGAAGAGACAGATTCAAGAGACAGTCCGCACTCCGGGAACATGACGAGTTGAA 348  
QY 61 LeuLeuLeuSerThrLeuGluIlyGlyValTPrhIleuGlyTPrhArgGluPheVal 80  
DB 349 CTGCTGCTGAGCACCTTGAGAAAGGAGTCTGGCACCTTGCTTGACTCGGGAATTCTGTG 408  
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
DB 409 GAGGCCCTCCGGAAACCGGACCCCTCTGCGCCGCTACATGAACCTGAGCTCAGC 468  
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
DB 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCAACTGCTGAACCTC 528  
QY 121 LeuGlnProThrLeuValAspIlyLeuLeuValArgAspValLeuAspIlyS CysMetGlu 140  
DB 529 CTTCAAGCCCACTCTGTGTGACAAGCTTCTAGTTAGAGACGCTCTTGATAGTCATGGAG 588  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
DB 589 GAGGAACCTGTTGACAAATTGAAGACAGAAACCGGATTGCTGTCAGAAACCAATGGAAT 648  
QY 161 GluSerGlyValArgGluLeuLeuIlyArgIleValGlnIlySgluAsnTrpPheSerAla 180  
DB 649 GAATCAAGGTGAAGAGAGCTAATAAAGGATTGTGAGAAAGAAAACGTGTTCTCTGCA 708  
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
DB 709 TTTCTGAATGNNNNNNNNNNNNNNNNNNNNATGAACCTTGCCAAGAGTTAACAGGCTCT 768  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220

DB 769 GATTGCTCAGAAAGCAATGACAGATTGAGAAATTATCACAAAGTTGATGCTCTCAAGTG 828  
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIlySgluValTrpGlyMet 240  
DB 829 GAGAAGCAACTCTTTTCAACACACAGTTCAGCCAAATCTGGAGAAGAGAGTCTGGGCATG 888  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
DB 889 GAGAAATACTCATCAGAAATCATCTTTTGAGATTCTTGTAGTTTCAGAAATCAGACACA 948  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
DB 949 AGTTTGACAGAAAGAGTGTACGCTGCTTAGATGAAAGTCTTGACATTAACAGCAACATG 1008  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
DB 1009 GGCAGTGATTCAGGCACCATGGGAGAGTGATTCAGATGAAGAAATGTGGCAAGCAAGCA 1068  
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
DB 1069 TCCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCACTTG 1128  
QY 321 GluGlyIlyAsnIleIleIleCysLeuProThrGlySerGlyIlySerThrArgValAlaVal 340  
DB 1129 GAAGGAAAGAAATATCATCATCTGCCCTCCCTACAGGAGTGGAAGAAACCAAGTGGCTGT 1188  
QY 341 TyrIleAlaIlyAspHisLeuAspIlySgluAspIlySgluAspGluProGlyIlyValIle 360  
DB 1189 TACATTGCCAAGAGATCATCTTAGACAAGAAAGAAAGCATCTGAGCCTGGAAGAAAGTTATA 1248  
QY 361 ValLeuValAsnIlySgluValLeuLeuValGluGlnLeuPheArgIlySgluPheGlnProPhe 380  
DB 1249 GTTCTTGCAATAAGGTACTGTAGTTGAACAGCTCTCCGCAAGAGTCCCAACCATTT 1308  
QY 381 LeuIlySgluSgluTyrArgValIleGlyLeuSerGlyAspThrGlnLeuIlySgluSerPhe 400  
DB 1309 TTGAAGAAATGTATCTGTATTGATTAGTGTGATACCCAACTGAAATATCATATT 1368  
QY 401 ProGluValValIlySgluSgluAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
DB 1369 CAGAAAGTGTCAAGTCTGTGATATTATATCATGATACAGCTCAAAATCTTGAAAACTCC 1428  
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
DB 1429 CTTCTTAAACTTGAAAAATGAGAAGATGTGTGTTCAATTGTACACTTTTCCCTCAT 1488  
QY 441 IleIleAspGluCysHisHisThrAsnIlySgluAlaValTyrAsnAsnIleMetArgHis 460  
DB 1489 ATCATGTGATGAATGTATCACACCAACAAGAGCAGTATATAATACATCATGAGCAT 1548  
QY 461 TyrLeuMetGlnIlySgluSgluAsnAsnArgLeuIlySgluAsnIlySgluProValIlePro 480  
DB 1549 TATTGTATGCAGAGTGTGAAAAACAATAGACTCAAGAAAGAAAACAACCATGATTCCC 1608  
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrIlySgluAla 500  
DB 1609 CTTCTCAGATACTGGGACTTAACAGCTTCACTGTGTGAGGGGCCACGAAAGCAAGCC 1668  
QY 501 IlySgluGluIlySgluIleLeuIlySgluCysAlaAsnLeuAspAlaPheThrIleIlySthr 520  
DB 1669 AAAGCTGAAGAACACACTTTTAAACTATGTGCAATCTTGATGATTAATTAATAAACT 1728  
QY 521 ValIlySgluAsnLeuAspGlnLeuIlySgluAsnGlnIleGlnIlyProCysIlySgluPheAla 540  
DB 1729 GTTAAAGAAAACCTGTATCAACTGAAAAACCAATAACAGAGCCATGCAAGATTGGCC 1788  
QY 541 IleAlaAspAlaThrArgGluAspProPheIlySgluIlySgluLeuGluIleMetThrArg 560  
DB 1789 ATTGCAGATGCACACGAGAGAGATCCATTTAAGAGAACTTCTAGAAATATGACAAAG 1848  
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580

Db 1849 ATTCAACTTATGTGCAAAATGAGTCCCAATGTCAGATTTTGGAACTCAACCCCTATGAACAA 1908  
Qy 581 TTPAlaIleGlnMetGluYsLysAlaAlaYsLysGlyAsnArgLysGluArgValCys 600  
Db 1909 TGGCCATTCAATGGAAGAAAAAGCTGCMAAGAGAGAAATCGCAAGAACGTGTTTGT 1968  
Qy 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
Db 1969 GCAGAACATTTGAGGAAGTCAATGAGGCCCTCAAAATTAATGACACAATTCGAATGATA 2028  
Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluYsAspLysLysPheAlaVal 640  
Db 2029 GATGCGTATACTCATCTGAAACTTTCTATAATGAGAGAAAGATAAGATTGCAAGTC 2088  
Qy 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 2089 ATAGAAAGATGATAGTAGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAGATGAG 2148  
Qy 661 AspAspLeuLysLysPProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 2149 GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATGATTTCTCATGACTTTATTT 2208  
Qy 681 PheGluAsnAsnLysMetLeuYsArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2209 TTTGAAAAACAATAAATGTTGAAAAAGCTGGCTGAAAAACCAAGATATGAAAAAG 2268  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720  
Db 2269 CTGACCAATTAAGAAATACCAATATGAGCAATATACTAGAGACTGAGGAATCAGACAGA 2328  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740  
Db 2329 GGAATTAATCTTTACAAAAACACGACAGAGTGCAATATGCCGCTTCCCACTGGAATTA 2388  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisLysLeuIleGlyAlaGlyHisSer 760  
Db 2389 AATGAAAAATTGTGCTGAGTAGAGGTCAAAGCCCACTCTGATTGGAAGCTGACACAGC 2448  
Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2449 AGTGAGTTCAAAACCATGACACAGATGAACAAAAAGAGTCATTAGTAAATTTCGCACT 2508  
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2509 GGAAAAATAAATCTGCTTATCGCTACCAAGTGCAGAGAAAGTCTGATATTAAAGAA 2568  
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2569 TGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAATAGCCATGTGCCAGCCCGT 2628  
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2629 GGTCCAGCCAGAGCTGATGAGACCACTACGTCCTGGTTCCTCACAGTGTTCCAGAGTT 2688  
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860  
Db 2689 ATCGAACGTGAGACAGTTAATGATTTCCGAGAGAGATGATGTAATAAGCTATACATTGT 2748  
Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2749 GTTCAAAATATGAAACAGAGAGATAGCTCATAAGATTTTGAATTACAGATGCAAAAGT 2808  
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900  
Db 2809 ATAAATGGAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTAACAAGATTAACCCA 2868  
Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920  
Db 2869 TCACATAATTACTTCTTTCGAAAAAAGTCAGAGTGCTAGCTGTTCTGGGGAAGATATC 2928  
Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
Db 2929 CATGTAATTGAGAAAAATGCATCACGTCATATATGACCCAGAAATTCAAGAACTTTACATT 2988

Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 2989 GTAAGAGAAAAACAAAGCACTGCAAAAAGAGTGTCCGACTATCAAAATTAATGCTGAATC 3048  
Qy 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 3049 ATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAAGATTGCCCT 3108  
Qy 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000  
Db 3109 TGTCCAAATATAGGAATTTTGTAGTGTTTCAAAAATATTAACAAGAAAGAAACAATAC 3168  
Qy 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020  
Db 3169 AAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAGAATGCTGTTTA 3228  
Qy 1021 PheSerAspGluAsp 1025  
Db 3229 TTTAGTGATGAGGAT 3243

RESULT 14

ADJ75813  
ID ADJ75813 standard; DNA; 3771 BP.

XX ADJ75813;

DT 20-MAY-2004 (first entry)

XX 20-MAY-2004 (first entry)

DE Marker gene SEQ ID NO:1065.

XX bronchial asthma; chronic obstructive pulmonary disease;

KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;

KW gene therapy; marker gene; gene; ds.

OS Mus musculus.

XX EP1394274-A2.

XX 03-MAR-2004.

PF 04-AUG-2003; 2003EP-00254857.

XX 06-AUG-2002; 2002JP-00229312.

PR 20-MAR-2003; 2003JP-00077212.

XX (GENO-) GENOX RES INC.

XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;

XX WPI; 2004-193155/19.

PS Claim 14; SEQ ID NO 1065; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma  
CC or chronic obstructive pulmonary disease. The method comprises  
CC determining the expression level of a marker gene in a biological sample  
CC from a subject, comparing the expression level determined with the  
CC expression level of the marker gene in a biological sample from a healthy  
CC subject, and judging whether the subject has bronchial asthma or chronic  
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of  
CC genes (S1) whose expression levels increase when respiratory epithelial  
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)  
CC whose expression levels decrease when respiratory epithelial cells are  
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for  
CC testing for bronchial asthma or chronic obstructive pulmonary disease;  
CC (2) a kit for screening for a candidate compound for a therapeutic agent  
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)

CC an animal model for bronchial asthma or chronic obstructive pulmonary  
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a  
CC method for producing an animal model for bronchial asthma or chronic  
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial  
CC asthma or chronic obstructive pulmonary disease, comprising the compound,  
CC a marker gene or an antisense nucleic acid corresponding to a portion of  
CC the marker gene, a ribozyme, a polynucleotide that suppresses the  
CC expression of the gene through an RNAi effect or an antibody recognising  
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for  
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a  
CC probe has been immobilised to assay a marker gene. (I) has respiratory  
CC and antiasthmatic activities, and can be used in gene therapy. The method  
CC is useful for testing for or screening for a therapeutic agent for  
CC bronchial asthma or chronic obstructive pulmonary disease. The present  
CC sequence is used in the exemplification of the present invention.

XX Sequence 3771 BP; 1183 A; 821 C; 866 G; 901 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3771
Score:	4194.00	Matches:	817
Percent Similarity:	86.9%	Conservative:	75
Best Local Similarity:	79.6%	Mismatches:	132
Query Match:	79.0%	Indels:	2
DB:	12	Gaps:	2

US-09-515-363C-2 (1-1025) x ADJ5813 (1-3771)

Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPhearGlyrLeuIleSerCysPheArgAla 20  
Db 235 ATGTCGATTGTCGTTCGACAGGACAGCTTCAGGAATCTCATCTTATTCTTCAGGCC 294  
Qy 21 ArgValIysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
Db 295 AGGCTGAAGAAATGATCATTCAGGTGAGCCAGTGTGACCACTCATCTTCTGTCTGCA 354  
Qy 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 355 GAAACCAAGAGCAGATTCTTAAGATCAACACCTGTGTAAACACAGCGGCGCAGAA 414  
Qy 61 LeuLeuLeuSerThrLeuGluIlySGlyValTPrhIleuGlyTTrpThrArgGluPheVal 80  
Db 415 CTGCTGCTGAGCACCTTGAGCAGGACAATGGCCCTTGCGATGACGCGCAGATGTTCTG 474  
Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
Db 475 GAGGCCCTAGAGCAGATGGCAATCCCTAGCCGCGCTATGTCAAAACCCACACTCACT 534  
Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 535 GATCTGCCCTCTCTCTCTGAGACTGCCCATGACGAGTGTCTCCACTTGCTGACCCCTC 594  
Qy 121 LeuGlnProThrLeuValAspIlySLeuLeuValArgAspValLeuAspIlySLeuMetGlu 140  
Db 595 CTCGAGCCCACTTGTGTGACAACTTGTATTAACGATGCTTGACACTTGCCTTCGAG 654  
Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAngGlyAsn 160  
Db 655 AAGGACTATTGACAGTGAAGACAGAATCGAATTCGCTGTCGAGAAACAGCGGGAAT 714  
Qy 161 GluSerGlyValArgGluLeuLeuIlySArgIleValGlnIlySGluAsnTrpPheSerAla 180  
Db 715 GAGTCAGGTGTAAAGAGAGCTGTGAGAAAGATTGTGAGAAAGAAACTGTTTCTAAC 774  
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAngGluLeuValGlnIlySGluLeuThrGlySer 200  
Db 775 TTCCTGATGTTCTGCGCAAACTGGAATGATGACATATTCGAAGAACTAACAGGTGGA 834  
Qy 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 835 GGCTGCCCAAGACAAACACAGACTTGCTACTCGTCTCACAGAGATGGGCTGCAAGCT 894  
Qy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIlySGluValTrpGlyMet 240

Db 895 AATGAGTGTCTTCTGCTGCTGTCGATGAGTCAAGTCTGGAGACAGAGGCTGAACGTA 954  
Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
Db 955 GACGACATATTACAGAGGCTTCTGTACAGATTGCTGTGACCAACAGATCAGACACA 1014  
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 1015 AGTTGGCAGAAAGAAAGTTCAGCTGCTTCGATGAAGTCTTGACATTAACAGCAACATG 1074  
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnVal---AlaAlaArg 299  
Db 1075 GGCAGGATTCAAGCACCATGGAAAGTGAATGATGAAGTGTCAACAGACAAAGAA 1134  
Qy 300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319  
Db 1135 GTATCCCCCGAGCCAGAACTGACAGCTCAGGCCCTTACCAATGGAAGTGGCCCAACGCT 1194  
Qy 320 LeuGluGlyIlySAsnIleIleIleCysLeuProThrGlySerGlyThrArgValAla 339  
Db 1195 CTAGATGGGAAGATATTATTATCTGCTCCCAAGGGAGTGGGAAACAGAGTGCT 1254  
Qy 340 ValTyrIleAlaIlySAspHisLeuAspIlySlySlySlyAlaSerGluProGlyIlySVal 359  
Db 1255 GTTTACATCACCAAGATCACTTAGACAAAGAACAGGCAATCTGAATCCGGGAAGTT 1314  
Qy 360 IleValLeuValAsnIlySValLeuLeuValGluGlnLeuPheArgGlyGluPheGlnPro 379  
Db 1315 ATCGTCTTGTCAATAAGTAATGTTAGCAGAACAACTTCCGAAAGATTCAACCCA 1374  
Qy 380 PheLeuIlySlySlyTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuIlySlySer 399  
Db 1375 TATTGAAGAATAGTGATCGAATATTATGATTAAGTGGCGATACCCAGCTGAAGAAATATCA 1434  
Qy 400 PheProGluValValIlySserCysAspIleIleIleSerThrAlaGlnIleuGluAsn 419  
Db 1435 TTTCAGAGAAGTGTCAAACTTTACGATGTTATTATACAGCACTGCTCAATCTTGAAGAAC 1494  
Qy 420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439  
Db 1495 TCCCTCTTAATCTGAGAGTGAGACGATGACGCTGTGACGCTGTACAGCTTCTCTCTC 1554  
Qy 440 IleIleIleAspGluCysHisIleThrAsnIlySGluAlaValTyrAsnAsnIleMetArg 459  
Db 1555 ATTATCATTTGATGAGTGCCATCACCAACAAGAGGAGCTATTAACAACATCATGAGA 1614  
Qy 460 HisTyrLeuMetGlnIlySLeuIlySAsnAsnArgLeuIlySlySGluAsnIlySProValIle 479  
Db 1615 CGATATTGGAAGCAGAGCTGAGAAACAATGACCTCAAGAAACAAACCAACAGCCATT 1674  
Qy 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyIlyAlaThrIlySGln 499  
Db 1675 CCCCTGCCGAGATACTGAGACTGACAGCTTCACTGCTGTGAGAGCAGCCAAAGAGCAG 1734  
Qy 500 AlaIlySAlaGluGluHisIleLeuIlySLeuCysAlaAsnLeuAspAlaPheThrIleIlyS 519  
Db 1735 TCTGAGGCTGAAAAACATATTTTAAATATATGTGCCAATCTTGATGCCCTTTACATTAA 1794  
Qy 520 ThrValIlySGluAsnLeuAspGlnLeuIlySAsnGlnIleGlnIlyProCysIlySlyPhe 539  
Db 1795 ACAGTGAAGAGAAATCTTGCTCAACTCAAAACCAACCAATAAAGAACCATGCAAGAAATTT 1854  
Qy 540 AlaIleAlaAspAlaThrArgGluAspProPheIlySGluIlySLeuLeuGluIleMetThr 559  
Db 1855 GTGATTGCTGATGACACAGAGAAATCCATTAAAGAGAACTTCTAGAAATTATGCA 1914  
Qy 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579  
Db 1915 AGCATTCAGACTTACTGCCAAAAAAGTCCCAATGTCAGATTTTGGAAACCAACATATAGAG 1974  
Qy 580 GlnTrpAlaIleGlnMetGluIlySlySlyAlaAlaIlySlySGlyAsnArgIlySGluArgVal 599

Db 1975 CAGTGGCCATTCAATGGAGAAAAAGCTGCTAAAGACGGAAATCGCAAGATCGCGTC 2034  
Qy 600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619  
Db 2035 TGTGCAGAACATTGGAGAAATAACAAGCCCTTACAAATCAACGACAGCATCCGAATG 2094  
Qy 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAla 639  
Db 2095 ATTGATGCATATAGCCACCTGAGGCATTCTACACTGATGAGAAAGAAAGAAAGTTCCGA 2154  
Qy 640 ValIleGluAspAspSerAspGluGlyLysAspAspGluTyrCysAspGlyAspGluAsp 659  
Db 2155 GTCCTC--AATGACAGCGACAAGATGATGACGAGGCCACAGTTGCAATGACCAACTT 2211  
Qy 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679  
Db 2212 AAGGCCGATGTAAAGAAATCTTGAACCTGACGAAACGGATGAATTTCTCATGAATTTG 2271  
Qy 680 PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu 699  
Db 2272 TTCTTTGATTAACAAGAAATGTTGAAAAAACTAGCTGAAAAAACCAAATATCGAGAATGAA 2331  
Qy 700 LysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAla 719  
Db 2332 AAATCATTTAAATTAGAAACACGATACTGGAACAAATTCACAAAGTCTGAGGAGTCTCC 2391  
Qy 720 ArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThr 739  
Db 2392 CGAGGAATTATTTTACAAAACACAGACAGACACCTTCCAGCATTTCCAGTGATCATG 2451  
Qy 740 GluAsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHis 759  
Db 2452 GAAATGCAAAAGTTTGGGAGAGTTGAGTCAAAAGCGCATCACCTGATGGCGGGGCGAC 2511  
Qy 760 SerSerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArg 779  
Db 2512 AGCAGTGAAGTCAAGCCCATGACTCAGACTGAACAAAAAGAGTCATTAGTAAATTTCCG 2571  
Qy 780 ThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys 799  
Db 2572 ACTGGCGAAATAAATCTGCTTATCGCTACGACGGTGGCAGAAAGCCTTGATATCAAA 2631  
Qy 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819  
Db 2632 GAGTGCATATATTGTTATTCGTTATGGCCTTGTACGAAAGAGATAGCCATGCTCCAGGCC 2691  
Qy 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839  
Db 2692 CGGGGTGAGCCAGAGCTGATGAAGACAGTATGTCTGTGTCCAGCAGTGGCTCAGGA 2751  
Qy 840 ValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859  
Db 2752 GTTACCGAACGGAGATTGTTAATGATTTCCGAGAGAGATGATGATATAAGCTATTAAAC 2811  
Qy 860 CysValGlnAsnMetLysProGluGluTyrAlaHisIleLysIleLeuGluLeuGlnMetGln 879  
Db 2812 CGTGTTCAAAAATGAAACCAAGAGATATGCACATTAAGATTTTGAATTGACAGTGCAA 2871  
Qy 880 SerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsn 899  
Db 2872 AGTATCTTGAAAAAGAAATGAAAGTCAAAAAGAGCATTTGCCAAAGCAATACAAAGACAAAT 2931  
Qy 900 ProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAsp 919  
Db 2932 CCATCGTTAATAACACTTCTCTCAAAAATTTAGCATTGCTGTCTGCTCGGAGAAAAAC 2991  
Qy 920 IleHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyr 939  
Db 2992 ATCCATGTCAATTGAGAGATGCATCATGTCAATATGACACCAAGATTCAAGGAGCTTAC 3051  
Qy 940 IleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlu 959  
Db 3052 ATTGTAAGAGAAAACAAAGCACTGCAAAAAGAAATTTGCTGATTATCAGACCAAATGAGAG 3111

Qy 960 IleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisIleGlyLeuAspLeu 979  
Db 3112 ATTATCTGCAAAGTGTGGCCAGGCTTGGGAAACAATGATGGTGCACAAAGTTTAGATTGG 3171  
Qy 980 ProCysLeuLysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGln 999  
Db 3172 CCTGTCTTAAATAAGCAATTTGTAGTCAATTTCAAAAATAACTCACGAAGAAACAG 3231  
Qy 1000 TyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019  
Db 3232 TACAGAAGTGGGTGGAATTGCCCTATCAGATTTCTGATCTTGACTACTCAGAATACTGC 3291  
Qy 1020 LeuPheSerAspGluAsp 1025  
Db 3292 TTGTATAGTGATGAAGAT 3309  
RESULT 15  
AAS40960  
ID AAS40960 standard; cDNA; 1967 BP.  
XX  
AC AAS40960;  
XX  
XX 17-DEC-2001 (first entry)  
DT  
XX  
DE cDNA encoding novel human enzyme polypeptide #176.  
XX  
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
KW ligase; hyperproliferative disorder; immunodeficiency disorder;  
KW autoimmune disorder; neurological disorder; metabolic disorder;  
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;  
KW anti arthritic; nephrotropic; anticoagulant; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200155301-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001239.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.

PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226868P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	01-DEC-2000;	2000US-0250160P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250391P.
PR	12-SEP-2000;	2000US-0231968P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234998P.	XX		
PR	26-SEP-2000;	2000US-0235484P.	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	2000US-0235834P.	XX		
PR	27-SEP-2000;	2000US-0235836P.	PI	Rosen CA, Barash SC, Ruben SM;	
PR	29-SEP-2000;	2000US-0236327P.	XX		
PR	29-SEP-2000;	2000US-0236367P.	DR	WPI; 2001-465566/50.	
PR	29-SEP-2000;	2000US-0236368P.	DR	P-PSDB; AAU23090.	
PR	29-SEP-2000;	2000US-0236369P.	XX		
PR	29-SEP-2000;	2000US-0236370P.	PT	Novel polypeptides and polynucleotides useful for diagnosing, preventing,	
PR	02-OCT-2000;	2000US-0236802P.	PT	treating neural, immune system, muscular, reproductive, pulmonary,	
PR	02-OCT-2000;	2000US-0237037P.	PT	cardiovascular, renal, proliferative disorders and cancerous diseases.	
PR	02-OCT-2000;	2000US-0237038P.	XX		
PR	02-OCT-2000;	2000US-0237039P.	PS	Claim 4; SEQ ID NO 186; 1180pp; English.	
PR	02-OCT-2000;	2000US-0237040P.	XX		
PR	13-OCT-2000;	2000US-0239935P.	CC	The present invention relates to the isolation of novel human enzyme	
PR	13-OCT-2000;	2000US-0239937P.	CC	polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences	
PR	20-OCT-2000;	2000US-0240960P.	CC	encoding them. The enzyme polypeptides of the invention may comprise the	
PR	20-OCT-2000;	2000US-0241221P.	CC	functional classes of oxidoreductases, transferases, hydrolases, lyases,	
PR	20-OCT-2000;	2000US-0241785P.	CC	isomerases or ligases. The sequences of the invention are useful in the	
PR	20-OCT-2000;	2000US-0241786P.	CC	diagnosis, treatment, prevention and/or prognosis of a wide range of	
PR	20-OCT-2000;	2000US-0241808P.	CC	disorders including hyperproliferative disorders (e.g. cancer),	
PR	20-OCT-2000;	2000US-0241809P.	CC	immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.	
PR	20-OCT-2000;	2000US-0241826P.	CC	arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic	
PR	01-NOV-2000;	2000US-0244617P.	CC	disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),	
PR	08-NOV-2000;	2000US-0246474P.	CC	cardiovascular disorders (e.g. atherosclerosis), blood-related disorders	
PR	08-NOV-2000;	2000US-0246475P.	CC	(e.g. haemophilia), reproductive disorders (e.g. infertility) and	
PR	08-NOV-2000;	2000US-0246476P.	CC	infectious disorders (e.g. Influenza). The polynucleotides of the	
PR	08-NOV-2000;	2000US-0246477P.	CC	invention can also be used in gene therapy. AAS40785-AAS41684 represent	
PR	08-NOV-2000;	2000US-0246478P.	CC	cDNA sequences encoding for the novel human enzyme polypeptides of the	
PR	08-NOV-2000;	2000US-0246523P.	CC	invention. Note: The sequence data for this patent did not form part of	
PR	08-NOV-2000;	2000US-0246524P.	CC	the printed specification, but was obtained in electronic format directly	
PR	08-NOV-2000;	2000US-0246525P.	CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences	
PR	08-NOV-2000;	2000US-0246526P.	XX		
PR	08-NOV-2000;	2000US-0246527P.	XX		
PR	08-NOV-2000;	2000US-0246528P.	XX		
PR	08-NOV-2000;	2000US-0246532P.	XX		
PR	08-NOV-2000;	2000US-0246609P.	XX		
PR	08-NOV-2000;	2000US-0246610P.	XX		
PR	08-NOV-2000;	2000US-0246611P.	XX		
PR	08-NOV-2000;	2000US-0246613P.	XX		

Sequence 1967 BP; 723 A; 344 C; 416 G; 476 T; 0 U; 8 Other;

Alignment Scores: 5.66e-251 Length: 1967  
Pred. No.: 3176.00 Matches: 636  
Score: 3176.00  
Percent Similarity: 97.1% Conservative: 2  
Best Local Similarity: 96.8% Mismatches: 18  
Query Match: 59.8% Indels: 5

DB:	4	Gaps:	1
US-09-515-363C-2 (1-1025) x AAS40960 (1-1967)			
Qy	294	GLuAsnValAlaAlaArgAlaSerProGluProGluLeuGlnLeuArgProTyrGlnMet	313
Db	3	GAGATGTGGCAGCAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCGCTTAACCAATG	62
Qy	314	GLuValAlaGlnProAlaLeuGlnGlyLysAsnIleIleCysLeuProThrGlySer	333
Db	63	GAA GTT GCC CAG CCA GCT TGG AAG GGA AGA AT AT C AT CT G C C T C C C T A C A G G A G T	122
Qy	334	GLySerThrArgValAlaValTyrIleAlaLysAspHisLeuAspLysLysLysAla	353
Db	123	GGAAAAACCAAGAGTGGCTGTTACATGGCCAAGATCACTTAGACAAGAAAAAGCA	182
Qy	354	SerGluProGlyLysValIleValLeuValAsnLysValLeuLeuValGlnGlnLeuPhe	373
Db	183	TCTGAGCCTGGAAGTTATAGTTCTTGTCATTAAGGTACTGCTAGTTGAACAGCTCTTC	242
Qy	374	ArgLysGluPheGlnProPheLeuLysLysTyrTyrArgValIleGlyLeuSerGlyAsp	393
Db	243	CGCAAGAGTTCCAACCATTTTGAAGAAATGGTATCGTGTATTGGATTAACTGGTGAT	302
Qy	394	ThrGlnLeuLysIleSerPheProGluValValLysSerCysAspIleIleIleSerThr	413
Db	303	ACCCA CTG AAA AT AT C AT T T C C A G A A G T T G C A A G T C C T G T G AT AT T A T T A T C A G T A C A	362
Qy	414	AlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGln	433
Db	363	GCTCAATCCTTGAAAACTCCCTCTTAAACTTGAAAAATGGAAGATGCTGTGTTCAA	422
Qy	434	LeuSerAspPheSerLeuIleIleLeuAspGluCysHisThrAsnLysGluAlaVal	453
Db	423	TTGTGACACTTTTCCCTCATTTATCATGTGATGATGTCTACACCAACAAGAACAGTGTG	482
Qy	454	TyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLys	473
Db	483	TATATAATCATCATGAGGCATTATTGTGATGCAGAA GTTGA AAAACAATA GACTCAAGAAA	542
Qy	474	GluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyVal	493
Db	543	GAAAA CA A A C C A G T A T T C C C T T C C T C A G A T A C T G G A C T T A A C A G C T T C A C C T G T G T T	602
Qy	494	GlyGlyAlaThrLysGlnAlaLysAlaGlnGlnHisIleLeuLysLeuCysAlaAsnLeu	513
Db	603	GGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTT	662
Qy	514	AspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGln	533
Db	663	GATGCATTTACTATTAAAACTGTTTAAGAAAACTTGATCAACTGAAAAACCAATAACAG	722
Qy	534	GluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLys	553
Db	723	GAGCCATGCCAAGAGTTTGGCCATTGCCAGATGCACCAGAGAGATCCATTTAAAGAGAAA	782
Qy	554	LeuLeuGlnIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe	573
Db	783	CTTCTAGAATAATGACAAGATTCAAACTTATTGTCAATGAGTCCAAATGTCAGATTTT	842
Qy	574	GlyThrGlnProTyrGlnGlnTyrAlaIleGlnMetGluLysLysAlaAlaLysLysGly	593
Db	843	GGAAC TCA A C C C T T A T G A A C A A T G G C C A T T C A A A T G A A A A A A A A G C T G C A A A A G A A G A	902
Qy	594	AsnArgLysGluArgValCysAlaGlnHisLeuArgLysTyrAsnGluAlaLeuGlnIle	613
Db	903	AATCGCAAGACGTTGTTGTGCAGAACATTTGAGGAAGTAACAATGAGGCCCTACAAATT	962
Qy	614	AsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGlnuThrPheTyrAsnGlnGlu	633
Db	963	AATGACACAATTCCGAATGATAGATCGTATACTCATCTTGAAACTTTCTATAATGAAAGAG	1022
Qy	634	LysAspLysLysPheAlaValIleGluAspAspSerAspGlnGlyGlyAspAspGluTyr	653

Db	1023	AAAGATAAGAA GTT T G C A G T C A T A G A A G A T G A T A G T G A G G G T G T G A T G A G T A T	1082
Qy	654	CysAspGlyAspGluAspGluAspAspLeuLysLysProLeuLysLeuAspGluThrAsp	673
Db	1083	TGTGATGTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGAT	1142
Qy	674	ArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsn	693
Db	1143	AGATTTCTCATGACTTATTTTGTGAAAACAAATAAAATGTTGAAAAGCTGCGTGA AAAAC	1202
Qy	694	ProGluTyrGluAsnGlnLysLeuThrLysLeuArgAsnThrIleMetGlnGlnTyrThr	713
Db	1203	CCAGATATGAAATGAAAAAGCTGACCAAAATTAAGAATACCATATGAGCAATATACT	1262
Qy	714	ArgThrGlnGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAla	733
Db	1263	AGGACTGAGGAATCAGCAGAGGAATATCTTTACAAAACACGACAGAGTCATATGCG	1322
Qy	734	LeuSerGlnTyrIleThrGluAsnGlnLysPheAlaGluValGlyValLysAlaHisHis	753
Db	1323	CTTTCCCACTGATTTATGAAAATGAAAAATTTGCTGAAGTAGAGTCAAA GCCCACCAT	1382
Qy	754	LeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGlnGlnLysGlu	773
Db	1383	CTGATTTGAGCTGGACACAGCAGTGA GTTCAA ACCATGACACAGAA TGAACAAAAGAA	1442
Qy	774	ValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGlu	793
Db	1443	GTCA TTAGTAATTTCCGCACTGGA AAAATAAATCTGCTTATCGCTCC -ACAGTGGCAGAA	1501
Qy	794	GlnGlyLeuAspIleLysGlnCysAsnIleValIleArgTyrGlyLeuValThrAsnGlu	813
Db	1502	GAA G C T G T G A T A T T A A G A A T G T A C A T T G K T A T C C G T T A T G C T C A C C A T A T G A A	1561
Qy	814	IleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuVal	833
Db	1562	ATACC -ATG GTCCAGCC -CGTGTGAGCTAGAGCTGATGAGAGCCTAACGT -CTGGTT	1618
Qy	834	AlaHisSerGlySerGlyValIleGlnHisGluThrValAsnAspPheArgGluLysMet	853
Db	1619	GCTCACA GTG GTT CAGAGTTATCGAA C GTGAGACAGTTAATGATTTCCGAGAGAA GATG	1678
Qy	854	MetTyrLysAlaIleHisCysValGlnAsnMetLysProGlnGluTyrAlaHisLysIle	873
Db	1679	ATGTATAAAGCTATC ---ATTGKGTTCAAATATGAACCAAGAGATGCTYATAAGAWT	1735
Qy	874	LeuGlnLeuGlnMetGlnSerIleMetGlnLysLysMetLysThrLysArgAsnIleAla	893
Db	1736	TTGGAATTTACAGATGCMAAGTATATGGA AAAAGAAAATGAAAACCAAGACAATATTGCC	1795
Qy	894	LysHisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeu	913
Db	1796	AAGCATTAACAAGATAACCCATCACTAATTAACTTTCTTGCAAAAACCTGCAGTGTCTA	1855
Qy	914	AlaCysSerGlyGluAspIleHisValIleGlnLysMetHisHisValAsnMetThrPro	933
Db	1856	GCTGTCTCTGGGAAGATATCCATGNAATTGANVAAAATGCATNACGTCAATATGACCCCA	1915
Qy	934	GluPheLysGlnLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLys	950
Db	1916	GAATTCAAGAACTTTACTTTGGAAGAGAAAAACAAGCACTGC AAAAGAAA 1966	

Search completed: March 12, 2006, 22:52:31  
Job time : 2545 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 13, 2006, 15:21:08 ; Search time 9400 Seconds  
(without alignments)  
5101.783 Million cell updates/sec

Title: US-09-515-363C-2  
Perfect score: 5311  
Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNLDYSECLFSDSD 1025

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 segs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abs/ABSSWEB.spool/US09515363/runat\_10032006\_181130\_7973/app.query.fasta\_1  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs04  
-USER=US09515363 @CGN\_1\_1\_5315 @runat\_10032006\_181130\_7973 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hlc:\*  
5: gb\_hlc:\*  
6: gb\_est4:\*  
7: gb\_est5:\*  
8: gb\_est6:\*  
9: gb\_est7:\*  
10: gb\_gss1:\*  
11: gb\_gss2:\*  
11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5250	98.9	3078	11	DQ050960	DQ050960 Homo sapi
2	4879	91.9	3078	11	DQ050961	DQ050961 Pan trogl
3	2673	50.3	2304	4	AK037057	AK037057 Mus muscu
4	2633	49.6	2331	4	AK040519	AK040519 Mus muscu
5	1597	30.1	1013	3	BQ233683	BQ233683 AGENCOURT
6	1376	25.9	870	5	BQ960157	BQ960157 AGENCOURT
7	1366	25.7	2530	4	CR857114	CR857114 Pongo pyg

8	1289.5	24.3	918	5	BU189982	BU189982 AGENCOURT
9	1257	23.7	1174	5	BU902097	BU902097 AGENCOURT
10	1218	22.9	1115	3	BM467983	BM467983 AGENCOURT
11	1206.5	22.7	1035	2	BF337464	BF337464 602035195
12	1186	22.3	690	8	DR422160	DR422160 nav0805.
13	1129	21.3	729	2	BF983236	BF983236 602305873
14	1122	21.1	992	5	BY720783	BY720783 BY720783
15	1122	21.1	1197	4	AK018602	AK018602 Mus muscu
16	1098	20.7	827	2	BG741146	BG741146 602631817
17	1094.5	20.6	781	2	BF686405	BF686405 602143786
18	1083	20.4	653	7	CR752055	CR752055 DKFZp469M
19	1081.5	20.4	1041	2	BE882040	BE882040 601505326
20	1078	20.3	1046	3	BM476961	BM476961 AGENCOURT
21	1066	20.1	627	5	BX492926	BX492926 DKFZp781C
22	1059	19.9	755	5	BQ772836	BQ772836 UI-H-FE0-
23	1055.5	19.9	1239	3	BM467774	BM467774 AGENCOURT
24	1045	19.7	4098	4	BC070029	BC070029 Homo sapi
25	1021.5	19.2	800	7	CO396840	CO396840 AGENCOURT
26	1019	19.2	582	1	AU310289	AU310289 AU310289
27	1016	19.1	672	3	BQ316075	BQ316075 CM3-CT027
28	1016	19.1	672	3	BQ316108	BQ316108 CM3-CT027
29	1016	19.1	672	3	BQ316120	BQ316120 CM3-CT027
30	994	18.7	671	6	CB453859	CB453859 709752 MA
31	990.5	18.6	911	3	BI454996	BI454996 603173580
32	990	18.6	672	6	CA339255	CA339255 NISC_1x12
33	989.5	18.6	1611	3	BM455238	BM455238 AGENCOURT
34	969	18.2	882	8	DN081906	DN081906 JGI CABD1
35	954	18.0	682	6	CF363618	CF363618 833509 MA
36	952.5	17.9	932	2	BG967530	BG967530 602833306
37	949.5	17.9	585	3	BP210040	BP210040 BP210040
38	947	17.8	647	6	CB424267	CB424267 598511 MA
39	942	17.7	674	7	CV194473	CV194473 SNESTbabo
40	930.5	17.5	698	7	CO748417	CO748417 SNESTbaa6
41	928.5	17.5	701	6	CA423868	CA423868 UI-H-FE1-
42	911.5	17.2	568	3	BQ311714	BQ311714 QV3-BN004
43	903.5	17.0	755	1	AM038788	AM038788 AM038788
44	903.5	17.0	806	1	AM039330	AM039330 AM039330
45	900	16.9	2398	4	BC027369	BC027369 Mus muscu

ALIGNMENTS

RESULT 1	DQ050960	3078 bp	DNA	linear	GSS 02-JUN-2005
LOCUS	DQ050960				
DEFINITION	Homo sapiens MDAS gene, VIRTUAL TRANSCRIPT, partial sequence.				
ACCESSION	DQ050960				
VERSION	DQ050960.1	GI:66904159			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3078)				
AUTHORS	Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeille,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees				
JOURNAL	(er) Plos Biol. 3 (6), E170 (2005)				
PUBMED	15869325				
REFERENCE	2 (bases 1 to 3078)				
AUTHORS	Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civeille,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
JOURNAL					
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.				

FEATURES  
source Location/Qualifiers  
1..3078  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="2"  
<1..>3078  
/gene="MDA5"  
/locus\_tag="NC18981"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 length: 3078  
Score: 5250.00 matches: 1015  
Percent Similarity: 99.1% conservative: 1  
Best Local Similarity: 99.0% mismatches: 9  
Query Match: 98.9% indels: 0  
DB: 11 gaps: 0  
US-09-515-363C-2 (1-1025) x DQ050960 (1-3078)  
QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
DB 1 ATGTCGAATGGGTATTCACAGACGAGATTCCGCTATCTCATCTCGTCTTCAGGACC 60  
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
DB 61 AGGGTGAATAATGTACATCCAGCTGAGCCTGTGCTGACTACTGACCTTCTGCTGCA 120  
QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
DB 121 GAGGTGAAGAGACGATTTCAGAGGACAGTCGCCACCTCCGGAAACATGACGGCAGTTGAA 180  
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTTrpHisLeuGlyTTrpThrArgGluPheVal 80  
DB 181 CTGCTGCTGAGCACCTTGGAAGGAGTCTGGCACCTTGTTGACTCGGGAATTCTGTG 240  
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
DB 241 GAGGCCCTCCGGAGAACGGGACCCCTCTGGCCGCCCTACATGAACCTGAGCTCACG 300  
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
DB 301 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATATCTCCAACGTCTGAACCTC 360  
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140  
DB 361 CTTCA GCCCACTCTGTGGACAAGCTTCTAGTTAGAGCGTCTTGATTAAGTCATGAG 420  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
DB 421 GAGGAACCTGTTGACCAATTGAAGACAGAAACCGGATTGCTGTCAGAAACAATGGAAT 480  
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180  
DB 481 GAATCAGGTGTAAGAGAGCTACTAAAAAGATTGTGCAGAAAGAAAACTGTTCTCTGCA 540  
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
DB 541 TTTCTGAATGTTCTTCGTCAACAAGGAACAATGAATGTCTCCAAGATTAAACAGGCTCT 600  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
DB 601 GATTGCTCAGAAAGCAATGCAAGATTGAGAATTATTCACAAGTTGATGGTCTCAAGTG 660  
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTTrpGlyMet 240  
DB 661 GAAGAGCAACTTCTTCAACCAAGTTCCAGCCAAATCTGAGAGAGAGGCTGGGGCATG 720  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValIleSerGluSerAspThr 260  
DB 721 GAGATAACTCATCAGAAATCATCTTTTGAGATTCTTGTGAGTTCTGAAATCAGACACA 780

QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
DB 781 AGTTTGGCAGAGGAAGTGTACAGCTGCTTAGATGAAGCTTTGGACATAACAGCAACATG 840  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
DB 841 GCCAGTGATTCAGGACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGACA 900  
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320  
DB 901 TCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAGTTGCCAGCCAGCCTTG 960  
QY 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340  
DB 961 GAAGGAGAATATCATCTGCTGCTCCCTTACAGGAGTGAAGAAAAACAGAGTGTGTT 1020  
QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360  
DB 1021 TACATTGCCAAGATTCACCTTAGACAGAGAAAAAGCATCTGAGCTGGAAGTTATA 1080  
QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380  
DB 1081 GTTCTTGTCATAATAGGTACTGCTAGTTGAACAGCTCTCCGCAAGAGATTCCAACCATTT 1140  
QY 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400  
DB 1141 TTGAAGAAATGTATCGTGTATTAGATTAGTGTGATACCCAACTGAAAAATATCATTT 1200  
QY 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420  
DB 1201 CCAGAAAGTTGTCAAAGTCTGTGATATATTATTCAGTACAGCTCAATCTTGAAAACTCC 1260  
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440  
DB 1261 CTCTTAAACTTGGAATAAGAGAGATGCTGTGTTCATATGTGACACTTTCCCTCATTT 1320  
QY 441 IleIleAspGluCysHisIleThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460  
DB 1321 ATCATTTGATGATGTATTCACACCAACAAGACAGTGTATATATCATGAGGCAT 1380  
QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480  
DB 1381 TATTTGATGCAGAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCATGATTCCC 1440  
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAla 500  
DB 1441 CTTCCCTCAGATACGTGGACTAAACAGCTTCACTGTGTGGAGGGCCACGAAGCAAGCC 1500  
QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520  
DB 1501 AAAGCTGAAGAACACACTTTTAAAACTATGTGCCAATCTTGATGCAATTACTATTAACCT 1560  
QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540  
DB 1561 GTTAAAGAAAACTGTGATCAACTGAAAAACCAATAACAGAGCCATGCAAGAGTTGCC 1620  
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGlnIleMetThrArg 560  
DB 1621 ATTGCAGATGCAACCAAGAGATCCATTTAAGAGAACTTCTAGAAATTAATGACAAAGG 1680  
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580  
DB 1681 ATTCAAACTTATGTCAATAGAGTCCAATGTCAAGATTTTGAACCTCAACCTATGAACAA 1740  
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600  
DB 1741 TGGGCCATTCAAATGGAAGAAAAAGCTGCAAAAGAGAAATCGCAAGAAAGTGTGTGT 1800  
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620  
DB 1801 GCAGAACATTTGAGAGAGTACAAATGAGGCCCTTACAAATTAATGACACAAATTCGAATGATA 1860  
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAlaVal 640

Db 1861 GATGCGTACTCATCTTGAACCTTCTATATGAAGAGAAGATAAGAGCTTGCAATC 1920  
Qy 641 ILeGluAspAspSerAspGluGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660  
Db 1921 ATAGAGATGATAGTAGAGGGTGGTGTATGATGAGTATTGTGATGCTGATGAGAATGAG 1980  
Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680  
Db 1981 GATGATTTAAGAAACCTTTGAAACTGCATGAACACAGATGATTCTCATGATTTATT 2040  
Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700  
Db 2041 TTGAAAACAATAAATGTTGAAAAAGCTGGCTGAAACCAGAAATATGAAAAAG 2100  
Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720  
Db 2101 CTGACCAATTTAAGAAATACCAATAAGCAATATAGCACTGAGGAATCAGCAGCA 2160  
Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTripletThrGlu 740  
Db 2161 GGAATATCTTTACAAAAACAGACAGAGTGCATATGCGCTTCCAGTGGAATTAAGTGA 2220  
Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer 760  
Db 2221 AATGAAAAATTTGCTGAAGTAGAGTCAAAAGCCCACTCTGATTGAGCTGACACAGC 2280  
Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780  
Db 2281 AGTGANNNNNNNNNNNNNNNNNNNAATGAACAAAAAGAGTATTGTAATTTCCACT 2340  
Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800  
Db 2341 GGAAAAATAAATCTGCTTATCGCTAACACAGTGGCAGAGAAGTCTGATATTAAAAAG 2400  
Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820  
Db 2401 TGTAACATGTTATCCGTTATGCTCTCGTCAACCAATGAATAGCCATGCTCCAGCCCGT 2460  
Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840  
Db 2461 GGTGAGCCAGAGCTGATGAGACCACTACGTCCTGGTTGCTCACAGTGGTTCAAGAGTT 2520  
Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860  
Db 2521 ATGAACTGAGACAGTGAATGATTTCCGAGAGAAGATGATGATAAAGCTAATACATTGT 2580  
Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880  
Db 2581 GTTCAAAATATGAACCAAGAGATGCTCATAGATTTTGGATTAACAGATGCAAAAGT 2640  
Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnPro 900  
Db 2641 ATATGGAAGAAAGAAATGAAGCAAGAAATATGCCAAGCATTAACAAGATTAACCCA 2700  
Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920  
Db 2701 TCACTAATACTTCTCTTGCAAAAAGTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATC 2760  
Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940  
Db 2761 CATGTAATTGAGAAATGCAATCACTCAATATGACCCAGAAATTCAGGAATTTACATT 2820  
Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 2821 GTAGAGAAAAACAACACACTGCAAAAAGAGTGTGCCCACTATCAATAATAATGCTGAATC 2880  
Qy 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980  
Db 2881 ATCTGCAAAATGTGGCAGGCTGGGGAACAATGATGTGCACAAGAAGCTTAGATTGGCCT 2940  
Qy 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000

Db 2941 TGTCTCAAAATAGGAATTTGTAGTGTGTTTCAAAAATAATTCAACAAAGAAACAATAC 3000  
Qy 1001 LysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLeu 1020  
Db 3001 AAAAAGTGGTAGAATTAACCTATCACATTTCCCAATCTTGACTATTGAGAAATGCTTTA 3060  
Qy 1021 PheSerAspGluAsp 1025  
Db 3061 TTTAGTATGAGGAT 3075  
RESULT 2  
DQ050961 3078 bp DNA linear GSS 02-JUN-2005  
LOCUS DQ050961  
DEFINITION Pan troglodytes MDA5 gene, VIRtual TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION DQ050961  
VERSION DQ050961  
KEYWORDS DQ050961.1 GI:66904160  
SOURCE GSS.  
ORGANISM Pan troglodytes (chimpanzee)  
Pan troglodytes  
Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Pan.  
REFERENCE 1 (bases 1 to 3078)  
AUTHORS Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,  
Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeello, D.,  
White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.  
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and  
Chimpanzees  
JOURNAL (er) PLOS Biol. 3 (6), E170 (2005)  
PUBMED 15869325  
REFERENCE 2 (bases 1 to 3078)  
AUTHORS Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,  
Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeello, D.,  
White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.  
TITLE Direct Submission  
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment. Translation starts at the beginning of  
alignment.  
FEATURES  
source location/Qualifiers  
1..3078  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
gene <1..>3078  
/gene="MDA5"  
/locus\_tag="HC18981"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 3078  
Score: 4879.00 Matches: 952  
Percent Similarity: 93.1% Conservative: 2  
Best local Similarity: 92.9% Mismatches: 71  
Query Match: 91.9% Indels: 0  
DB: 11 Gaps: 0  
US-09-515-363C-2 (1-1025) x DQ050961 (1-3078)  
Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
Db 1 ATGTGCAATGGGTGTTCCACAGACGAGAATTTCCGCTATCTCATCTCGTCTTCAAGGCC 60  
Qy 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
Db 61 AGGGTGAATAATGTACATCCAGGTGAGCTGTGCTGGACTACCTGACCTTTCTGCTGCA 120  
Qy 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 121 GAGGTGAAGAGACAGATTCAAGAGACAGTGCACCTCCGGAAATGACGAGGAGTTGAA 180

QY	61	LeuLeuLeuSerThrLeuGluLysGlyValTyrPheIleGluGlyTyrThrArgGluPheVal	80
Db	181	CTGCTGCTGAGCACCTTGGAGAAAGGAGTCTGGCACCTTGGTGGACTCGGAATTCGTG	240
QY	81	GluaIaleuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
Db	241	GAGGCCCTCCGAGAACCGGCAGCCCTCTGCGCCGCGCTACATGAACCTGAGCTCACG	300
QY	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	301	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGATAATATCTCCAACCTGCTGAACCTC	360
QY	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
Db	361	CTTCAGCCCACTCTGGTGGACAACTTCTAGTTAGAGAGCTCTTGATAAGTGCAITGAG	420
QY	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	421	GAGGAAGTGTGACAAATGAGACAGAAACCGGATGTGCTGCAGAAAAACAATGGAAT	480
QY	161	GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTyrPheSerAla	180
Db	481	GAGTCAGGTGTAAAGAGAGCTACTAAAAAGGATTGTGCAGAAAAAGAACTGTTCTCTGCA	540
QY	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	541	TTTCTGAATGTTCTTCGTCAAACAGGAAACAATGAACCTTGTCCAAGAGTTAACAGGCTCT	600
QY	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	601	GATTGCTCAGAAAGCAATGCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	660
QY	221	GluGluGlnLeuLeuSerThrValGlnProAsnLeuGluLysGluValTyrGlyMet	240
Db	661	NN	720
QY	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260
Db	721	NN	780
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
Db	781	AGTTGGCAGAAAGAGTGTCACTGCTTAGATGAAAGTCTTGACATTAACAGCAATG	840
QY	281	GlySerAspSerGlyThrMetCylSerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	841	GGCAGTGATTCAGGCACCATGGGAAGTATTGATGAGAGAAATGTGGCAGCAAGACGA	900
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetCylValAlaGlnProAlaLeu	320
Db	901	TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAAAATGGAAGTTGCCAGCCAGCCTTG	960
QY	321	GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
Db	961	GAAAGGAAGAATATCATCATCTGCTCCCTCACAGGAGTGGAAGAAAAACAGAGTGGCTGTT	1020
QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle	360
Db	1021	TACATTGCCCAAGGATCACTTAGACAAGAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATA	1080
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380
Db	1081	GTTCTTGTCAATAAAGTACTGTAAGTTGAACAGCTCTTCCGCAAGGAGTTCACCAACCATTT	1140
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1141	TTGAAGAAATGATCGTGTATTGGAATTAAGTGTGATACCCAATGAAATATATCATTT	1200
QY	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1201	CCAGAAGTGTCAAGTCCGTGTATATATATATATACAGTACAGCTCAAAATCCTTGAAAACTCC	1260
QY	421	LeuLeuAsnLeuGluAsnGlyCylAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440

Db	1261	CTCTTAACTTGAAAAATGGAGAAGATGCTGGTGTTCATTGTTCAGACTTTTCCCTCATTT	1320
QY	441	ILLEIleAspGluCySHisHisIsthrrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1321	ATCATTTGATGATGTGCATGCACCAACAAGAGACAGTGTATAATAACATCATGAGGCGT	1380
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
Db	1381	TATTTCATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCACTGATTCCC	1440
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAla	500
Db	1441	CTTCCTCAGATACTGGAGACTAACAGCTTCACCTGGTGTGGAGGGCCACGAAGCAAGCC	1500
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1501	AAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATGCATTACTTAAAAACT	1560
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	1561	GTTAAAGAAAACTTGATCAACTGAAAAAACCAATACAGAGCCATGCAGAAAGTTTGCC	1620
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
Db	1621	ATTGCAGATGCACACAGAGAGATCCATTTAAAGAGAAACTNNNAGAAATATGACAAAG	1680
QY	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	1681	ATTCAAACCTTATGTCAAAATAGATCCAAATGTCAAGATTTTGGAACTCAACCCATGAACAA	1740
QY	581	TyrAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1741	TGGGCCCATTCAAATGGAAAAAAAGCTGCAAAAGAAAGAAATCGCAAGACGTGTTGT	1800
QY	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	1801	GCAGACATTTGAGGAGTCAATGAGGCCCTACAAATTAATGACACTATTCCGAATGATA	1860
QY	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal	640
Db	1861	GATGCGTACTCATCTTGAAACTTTCTATATGAGAGAAAGATAGAAGTTTGCAATGC	1920
QY	641	IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	1921	ATAGACATGATAGTAGTAGGCGTGTGATGATGACTATTGTGATGCGATGAAGATGAG	1980
QY	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	1981	GATGATTTAAAGAAACCTTTGAAACTGGATGAACAGATAGATTCTCATGACTTTATTT	2040
QY	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db	2041	TTTGAAGAAACAATGAATGTTGAAAAAGCGTGGTGAAGAAACCAGAAATATGAAAAAG	2100
QY	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGlnSerAlaArg	720
Db	2101	CTGACCAAAATTAAGAAATACCAATAATGAGCAATATACTAGGACTGAGGAATCAGACGA	2160
QY	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGlu	740
Db	2161	GGAATATCTTTACAAAAACAGACAGAGTGCATATGCGCTTCCACAGTGAATTACTGAA	2220
QY	741	AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer	760
Db	2221	AATGAAGAAATTTGCTGAAGTAGAGTCAAGACCCACGNNNNNNNNNNNNNNNNNNNN	2280
QY	761	SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr	780
Db	2281	NNNNNNNNNNNNNNNNNNNNNNNNNAATGACACAAAAAGAAAGTCAATTAATTTCCGACT	2340
QY	781	GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu	800

Db 2341 GGAAAAATAATCTGCTTATCGCTACACAGTGGCAGAGAAGGCTCTGATATTAAAGAA 2400

Qy 801 CysAnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820

Db 2401 TGTACATTTGTTATCCGTATTGCTCTCGTACCCAATGAATAAGCATGTTNNNCCCGT 2460

Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValIleValAlaHisSerGlySerGlyVal 840

Db 2461 GGTGAGCCAGAGCTGATGAGACCACTACGTCCTGGTTGCTCACAGCGGTTCAGAGATT 2520

Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860

Db 2521 ATCGAAGTGAGACAGTAAATGATTTCCGAGAGAAGATGATGATAAGCTATACATTGT 2580

Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880

Db 2581 GTTCAAAATATGAACACAGAGAGTATGCTCATTAAGATTTTGAATTACAGATGCAAAAGT 2640

Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900

Db 2641 ATTAATGAAAAAGAAATGAAGAACCAAGAGAATATGCGCAAGCATTACAAAGATAACCCA 2700

Qy 901 SerLeuIleThrPheLeuCysLysAsnCysSerValIleuAlaCysSerGlyLysAspIle 920

Db 2701 TCACTAATACTTCTTCGCAAAACTGCAGTGTGCTAGCCGTGTTGGGGAAGATATC 2760

Qy 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940

Db 2761 CACGTAATTGAGAAATGCATCACTCATATGACCCAGAATTCAAGGAACCTTTACATT 2820

Qy 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960

Db 2821 GTAAGAGAAAAACAAGCACTGCAGAAAGAGTGTGCCCACTATCAATAAATGCTGAATC 2880

Qy 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980

Db 2881 ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTGGCCT 2940

Qy 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000

Db 2941 TGTCTCAAAATAAGGAATTTGTAGTGTTCCTTCAAAATAATTCACAAAGAAACAATAC 3000

Qy 1001 LysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGlyCysLysLeu 1020

Db 3001 AAAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAAGATGCTGTTTA 3060

Qy 1021 PheSerAspGluAsp 1025

Db 3061 TTTAGTGATGAGGAT 3075

RESULT 3

AK037057

LOCUS 2304 bp mRNA linear HTC 03-APR-2004

DEFINITION Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:9930105B04 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert sequence.

ACCESSION AK037057

VERSION AK037057.1 GI:26331913

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Komno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL 4

PUBMED 11076861

REFERENCE The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 2304)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Komno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

Direct Submission

TITLE Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

PLEASE VISIT OUR WEB SITE FOR FURTHER DETAILS.

URL:http://genome.gsc.riken.jp/

URL:http://fantom.gsc.riken.jp/.

FEATURES

source location/Qualifiers

1. .2304

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="FANTOM DB:9930105B04"

/db\_xref="taxon:10090"

/clone="9930105B04"

/sex="female"

/tissue\_type="vagina"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="adult"

224. .>2302

/note="unamed protein product; putative similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens] (SPTR|Q9H3G6, evidence: FASTY, 75.6%ID, 67.6%length, match=2079)"

CDS

```
/codon_start=1
/protein_id="BAC29687.1"
/db_xref="GI:26331914"
/translation="MSIVCSAEDSPFNILFRPRLKMYIQVEPVL DHLIFLSAETKE
OILKINTCGNTSAEELLSTLEQGMPLGWTOMFVEALEHSGNPLAARYKPTLDL
PSPSETAHDECLHLTLQPTLVKLLINDVLDTCPEKGLTVEDRNRIISAAGSNGN
ESGVRELRLRIQKENMFSTFLDVLROTNDALFOELTGGCPEDNTDLANS SHDGP
AANECLPAVIDESSLETEAMVNDILPEASCTDSVTTESDPSLAEGSVSCFDESIGH
NSNMGDSGTMGSDSDSESVIQTKRVSPEPELQLRPYQMEVAQPALDGKNIICLPFGS
GKTRAVYITKDHLDKKQASSESGYIVLVNKVMALEQLFRKFNPYLKMYRIIGLS
GDTOLKISPEVVKSYDVIISTAQILENSLNLNESGDDGVQLSDFSLIIDECHTN
KEAVYNIMRWRYLKQKLKRNNDLKKQKPAIPLPQILGLTASPGVGAAKQSEAEKHIL
NICANLDAFTIKTVKENLGOLKHQIKPECKFVIADDTRENPFKEKLEIMASIQTYC
OKSPMSDFGTQHYEQWAIQMEKKAADGNRKDRVCAEHLRKYNALQINDTIRMIDAY
SHLETFYTDKEKKFAVLNDSKSDDEASSCNDQLKGDVKS LKLDETDEFLMNLFED
NKKMLKKLAEN"
```

ORIGIN

Alignment Scores:

Pred. No.:	3,49e-275	Length:	2304
Score:	2673.00	Matches:	525
Percent Similarity:	84.3%	Conservative:	60
Best Local Similarity:	75.6%	Mismatches:	107
Query Match:	50.3%	Indels:	2
DB:	4	Gaps:	2

US-09-515-363C-2 (1-1025) x AK037057 (1-2304)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPhaArgTyrLeuIleSerCysPheArgAla 20  
Db 224 ATGTCGATGTGTCGTCGACAGACAGACGCTCAGGAATCTCATCTTATTCTTCAGGCC 283  
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
Db 284 AGGCTGAAATGTACATTCAGGTGAGCCAGCTGCTGACCACTCATCTTCTGCTGCA 343  
QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 344 GAAACCAAGAGCAGATTCTTAAAAAGATCAACACCTGTGTTAACACCCAGCGCGCAGAA 403  
QY 61 LeuLeuLeuSerThrLeuGlnLysGlyValTPrHisLeuGlyTyrThrArgGluPheVal 80  
Db 404 CTGCTGCTGAGCACCTTGAGACAGGACAAATGGCTCTGGGATGAGCGCAGATGTTCTG 463  
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
Db 464 GAGGCCCTAGAGCACAGTGGCAATCCCTAGCCGCGCTATGTCAAAACCACACTCACT 523  
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 524 GATCTGCCCTCTCCTCTCTGAGACTGCCCATGACGAGTGTCTCCACTTGTGACCCCTC 583  
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140  
Db 584 CTCACGCCCACTTTGGTGGACAAACTCTTGATTAACGATGTCTTGACACTTGTTCGAG 643  
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
Db 644 AAGGACTATTGACAGTCGAAGACAGAATAATCGGATTCTGCTGCAGAAACAGCGGAAT 703  
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180  
Db 704 GAGTCAGGTGTAAAGAGAGCTGCTGAAGAAGATTGTGCAGAAGAAAACGTGTTTCTACC 763  
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 764 TTCCTGATGTTCTCGCGCAACTGGAATGATGACTATTCCAGAACAATAACAGGTGGA 823  
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 824 GGCTGCCCAAGAACACACACACTTGGCTAACCTGCTCACAGAGATGGGCTTCAGACT 883  
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240

Db 884 AATGAGTGTCTTCTGCTGCTGCTGCTGATGAGTCAAGTCTGGAGACAGAGCCCTGGAACGTA 943  
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
Db 944 GACGACATATTACCAGAGGCTTCTGTACAGATTGCTGTGTGACCACAGAAATCAGACACA 1003  
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 1004 AGTTTGCAGAAGAGTGTACGCTGCTTCGATGAAGAAGCTTGGACATAACAGCAACATG 1063  
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnVal--AlaAlaArg 299  
Db 1064 GGCAGGATTTCAGGACCATGGAGAGTGAATTCAAGTGAAGTGTCTCCAGACAAAGA 1123  
QY 300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319  
Db 1124 GTATCCCCCGAGCCAGAACTGCAGCTCAGGCCCTTACCMAATGGAAGTGGCCCAACGACT 1183  
QY 320 LeuGluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAla 339  
Db 1184 CTAGATGGGAAGATATTATTATCTGCTCCCAACGGGAGTGGGAAAAACAGAGTGGCT 1243  
QY 340 ValTyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysVal 359  
Db 1244 GTTTACATCACCAAGATCACTTAGACAAGAAGAAGCAGCATCTGAATCCGGGAAGTT 1303  
QY 360 IleValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnPro 379  
Db 1304 ATCGTTCTTGCAATTAAGTAATGTTAGCAGAACAACTTTCCGAAAAGAGTTCACACCA 1363  
QY 380 PheLeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSer 399  
Db 1364 TATTTGAAGAAATGGTATCGAATTATTTGATTAAAGTGGCATACCCAGCTGAATAATATCA 1423  
QY 400 PheProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsn 419  
Db 1424 TTTCAGAAAGTTGTCAATCTTACGATGTTATTATCAGCACTGCTCAAAATCCTGAAAC 1483  
QY 420 SerLeuLeuAsnLeuGluAsnGlyLysAlaAspAlaGlyValGlnLeuSerAspPheSerLeu 439  
Db 1484 TCCCTCTTAATCTGGAGAGTGAGACGATGACGGTGTGACGCTGACACTTCTCTCTC 1543  
QY 440 IleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArg 459  
Db 1544 ATTATCATGTAGTGGCCATCACACCAAGAGGAGGAGTCTATACCAACATCATGAGA 1603  
QY 460 HisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIle 479  
Db 1604 CGATATTGGAAGCAGAACTGAGAAACAATGACCTCAAGAAACAACAACACAGCCATT 1663  
QY 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGln 499  
Db 1664 CCCCTGCCGACATACTAGGACTGACAGCTTCACCTGGTGTGAGCAGCAAAAAAGCAG 1723  
QY 500 AlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519  
Db 1724 TCTGAGGCTGAAAAACATATTTAAATATATGTGCCAATCTTGATGCTTTAACCATTA 1783  
QY 520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539  
Db 1784 ACAGTGAAGAGAACTTGGTCAACTCAACACCAAAATAAAGAACCATGCAAGAAATTT 1843  
QY 540 AlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThr 559  
Db 1844 GTGATTGCTGATGACACCAAGAAAATCCATTAAAGAGAACTTCTAGAAATTATGCA 1903  
QY 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579  
Db 1904 AGCATTCAGACTTACTGCCAAAAAGTCCAAATGTCAGATTTTGAACCAACATATATGAG 1963  
QY 580 GlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgVal 599

Db 1964 CAGTGGCCATTCAATGAGAAAAAGCTGTAAGAAGCGGAATCGCAAGATCGCGTC 2023

Qy 600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619

Db 2024 TGTGCAGAACATTGTGAGAGTACACAGAACCCCTTACAATCAACGACACGATCCGAATG 2083

Qy 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAla 639

Db 2084 ATTGATGCATATAGCCACTGAGACATTTCTACACTGATGAGAAAGAAAGAGTTGCGCA 2143

Qy 640 ValIleGluAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659

Db 2144 GTCCTC--AATGACAGCGACACAGAGTGATGACGAGCCACAGAGTGCATGACCAACTT 2200

Qy 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679

Db 2201 AAGGGCGATGTAAGAAATCTTTGAAACTGGACGAAACGGATGAATTTCATGAATTGG 2260

Qy 680 PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsn 693

Db 2261 TTCTTTGATACACAGAAATGTTGAAAAAACTAGCTGAAAAAC 2302

RESULT 4

AK040519 2331 bp mRNA linear HTC 03-APR-2004

LOCUS DEFINITION AK040519 2331 bp mRNA linear HTC 03-APR-2004

Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:A430105A06 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert sequence.

ACCESSION AK040519 GI:26333792

VERSION AK040519.1 GI:26333792

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 10349636

REFERENCE 2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 11042159

JOURNAL PUBMED 11042159

REFERENCE AUTHORS 3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Komno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillarary sequencer Genome Res. 10 (11), 1757-1771 (2000) 11076861

JOURNAL PUBMED 11076861

REFERENCE AUTHORS 4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 2331)

JOURNAL REFERENCE

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Komno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL:http://genome.gsc.riken.jp/

URL:http://fantom.gsc.riken.jp/.

location/Qualifiers

1..2331

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="FANTOM DB:A430105A06"

/db\_xref="taxon:10090"

/clone="A430105A06"

/tissue\_type="thymus"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="0 day neonate"

283..>2331

/note="unnamed protein product; putative similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens] (SPTR|QH3G6, evidence: FASTY, 75.6%ID, 67.6%length, match=2079)"

/codon\_start=1

/protein\_id="BAC30614.1"

/db\_xref="GI:26333793"

/translation="MSIVCSAEDSFRNLILFRRPLKMYIOVEPVLDDLIFLSAETKE QILKINTGNTSAEBLLSTLEGQMPLGWTQMFVEALEHSGNPLARYVKPTLTDL PPSSETADECLHLTLLOPTLVKLLINDVLDFCEKGLTVEDRNRISAAGNSGN ESGVRELRLRIVQENWFSTFLDYLRLQTDNALFQELTGGGCEPDNTDLANSHRDGP AANECLPAYDESSLTEAMNVDILPEASCTDSVTTESPTSLAEGSVSCFDESIGH NSNMGRDSGTMGSDSDSVIQTKEYSPEPELQRLPYQMEVAQPALDGKNIICLPTGS GTRVAVYITTKDHLDDKKQASESGYIVLVNKVMLAEQLFRKEFNPLYKKWYRIIGLS GDTOLKISPEPEVKSVDYDIISTAQILENSLNLNLESGDDGVQSDPSLIIDECHHTN KEAVYNNIMRYLKKQLRNNDLKKONKPAIPLPQILGLTASPGVGAAKKQSEAEKHIL NICANLDAFTIKTVKENIAGOLKQIKPECKKFVIADPTRENPFEKLLLEIMASIQTYC QKSPMSDFGTQHYEQWAIQMEKKAADGNRKDRVCAEHLRKYNELQINDTIRMIDAY SHLEFTYDEKEKKFAVLNDSKSDDEASSCNDQLKGDVYKSLKLDDETDFLMNLFPD N"

ORIGIN

Alignment Scores:

Pred. No.: 7.06e-271 Length: 2331

Score: 2633.00 Matches: 517

Percent Similarity: 84.2% Conservative: 59

Best Local Similarity: 75.6% Mismatches: 106

Query Match: 49.6% Indels: 2

DB: 4 Gaps: 2

US-09-515-363C-2 (1-1025) x AK040519 (1-2331)

Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20

Db 283 ATGTCGATGTCTGTTGCGACAGAGACAGCTTCAGGAATCTCATCTTATTCTTCAGGCC 342  
Qy 21 ArgValIysMetTyrIleGlnValGluProValIleuAspTyrLeuThrPheLeuProAla 40  
Db 343 AGCGTAAATATGATCATTCAGGTGAGCCAGCTGCTGACCACTCATCTTCTGTCTGCA 402  
Qy 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 403 GAAACCAAGACAGATCTTAAAGATCAACACCTGTGTAACACCGCGGACAGAA 462  
Qy 61 LeuLeuLeuSerThrLeuGluIlybGlyValTTrpHisLeuGlyTTrpThrArgGluPheVal 80  
Db 463 CTGCTGCTGAGCACCTTGAGCAGGACAATGGCCTCTGGGATGAGCGCAGATGTTCTGTG 522  
Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaIaArgTyrMetAsnProGluLeuThr 100  
Db 523 GAGGCCCTAGACACAGTGGCAATCCCTAGCCGCGCTATGTCMAAACCACTCACTCACT 582  
Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 583 GATCTGCCCTCTCCTTCCCTGAGAGACTGCCCATGAGAGTGTCTCCACTTGCTGACCCCTC 642  
Qy 121 LeuGlnProThrLeuValAspIlybLeuLeuValArgAspValIleuAspIySecMetGlu 140  
Db 643 CTCACGCCCACTTGGTGACAACTTCTGATTACGATGCTCTTGACACTTGCTTCGAG 702  
Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaIaGluAsnAsnGlyAsn 160  
Db 703 AAGGACTATTGACAGTCGAGACAGAAATCGGATTCTGCTGCAGAAACAGCGGGAAT 762  
Qy 161 GluSerGlyValArgGluLeuLeuLeuIyAsnArgIleValGlnIySGluAsnTrpPheSerAla 180  
Db 763 GAGTCAGGTGTAAAGAGAGCTGCTGAGAAGGATGTGCAAGAAAGAAAATGTTTCTACC 822  
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 823 TTCTCGATGTCTGCGCCAACTGGAATGATGCACTATTCGAAGAACTAACAGGTGA 882  
Qy 201 AspCysSerGluSerAsnIaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 883 GCGTGCCCAAGAACACACAGACTTGGCTAACCTCGTCTCACAGAGATGGCGCTGCAGCT 942  
Qy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluIySGluValTTrpGlyMet 240  
Db 943 AATGAGTGTCTTCTGCTGCTGCTGATGAGTCAAGTCTGAGACAGAGCGCCTGGAACGTA 1002  
Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260  
Db 1003 GACGACATATTACCAGAGGCTTCTGTACAGATTGCTGTGACCACAGAAATCAGACACA 1062  
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280  
Db 1063 AGTTTGGCAGAAAGAGTGTCACTGCTTCGATGAAGTCTTGACATTAACGCAACATG 1122  
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnVal--AlaAlaArg 299  
Db 1123 GGCAGGGATTCAAGCACCATGGGAAGTGTATCAGATGAAGTGTCAATCCAGACAAAAGA 1182  
Qy 300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319  
Db 1183 GTATCCCCCGAGCCAGAACTGCACTCAGGCCCTTACCAAAATGGAAGTGGCCCAACGACT 1242  
Qy 320 LeuGluGlyIyAsnIleIleCysLeuProThrGlySerGlyIySThrArgValAla 339  
Db 1243 CTAGATGGGAAGATATTATTCTGCTCCCAACGGGAGTGGGAAAAACAGAGTGGCT 1302  
Qy 340 ValTyrIleAlaIyAspHisIleuAspIySGlyIySGlyAlaSerGluProGlyIySGlyVal 359  
Db 1303 GTTATCATCACCAAGATCACTTAGACAAAGAAAGCAGCATCTGAATCCGGGAAGGTT 1362  
Qy 360 IleValIleuValAsnIySGlyLeuLeuValGluGlnPheArgIySGluPheGlnPro 379

Db 1363 ATCGTCTGTGCAATAGGTAATGTTAGCAGAACAACTTTCCGAAAAGAGTTCAACCA 1422  
Qy 380 PheLeuIySGlyTTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuIySileser 399  
Db 1423 TATTGGAAGAAATGTATTCGAATTATTGGATTAAAGTGGCATACCCAGCTGAAATATCA 1482  
Qy 400 PheProGluValValIySerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsn 419  
Db 1483 TTTCAGAAAGTTGTCAAAATCTTACGATGTTATTATCAGCACTGCTCAAAATCCTGAAAC 1542  
Qy 420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439  
Db 1543 TCCCTCTTAATCTGAGAGTGGAGACGATGACGGTGTGCAGCTGTGACACTTCTCTCTC 1602  
Qy 440 IleIleIleAspGluCysHisIleThrAsnIySGluAlaValTyrAsnAsnIleMetArg 459  
Db 1603 ATTATCATTTGATGAGTGCATCACACCAACAGAGGAGGTCTATTACACATCATGAGA 1662  
Qy 460 HisTyrLeuMetGlnIySGlyLeuIySGlyAsnAsnArgLeuIySGlyAsnIySGlyProValIle 479  
Db 1663 CGATATTTTGAAGCAGAGCTGAGAAACAATGACCTCAAGAAACAAACACAGCCATT 1722  
Qy 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyValAlaThrIySGln 499  
Db 1723 CCCCTGCCGAGATACTAGGACTGACAGCTTCACTGTGTGGAGCAGCCAAAAAGCAG 1782  
Qy 500 AlaIySGluGluHisIleLeuIySGlyCysAlaAsnLeuAspAlaPheThrIleIyS 519  
Db 1783 TCTGAGGCTGAAAAACATATTTAAATATATATGCCCCAATCTTGATGCCTTTACCATTA 1842  
Qy 520 ThrValIySGluAsnLeuAspGlnLeuIySGluAsnGlnIleGlnIyProCysIySGlyPhe 539  
Db 1843 ACAGTAAAGAGAAATCTTGGTCAACTCAACCAACCAATTAAGAAACCAATGCAAAATTT 1902  
Qy 540 AlaIleAlaAspAlaThrArgGluAspProPheIySGluIySGlyLeuLeuIleMetThr 559  
Db 1903 GTGATTGCTGATGACACACGAGAAATCCATTTAAAGAGAACTTCTGAATATGCA 1962  
Qy 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579  
Db 1963 AGCATTCAGACTTACTGCCCCAAAAAGTCCAAATGTCAGATTTTGGAAACCCCAATTATGAG 2022  
Qy 580 GlnTTrpAlaIleGlnMetGluIySGlyAlaAlaIySGlyAsnArgIySGluArgVal 599  
Db 2023 CAGTGGGCCATTCAAAATGAGAAAAAGCTGCTTAAAGACGGAATCGCAAGATCGCGCTC 2082  
Qy 600 CysAlaGluHisIleuArgIySGlyTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619  
Db 2083 TGTGAGAAACATTTGAGGAAGTAAACAGAACCCCTTCAAAATCAACGACAGATCCGAATG 2142  
Qy 620 IleAspAlaTyrThrHisIleuGluThrPheTyrAsnGluGluIySGlyAspIySGlyPheAla 639  
Db 2143 ATTGATGCATATAGCACCTGGAGACATTTCTACACTGATGAGAAAGAAAGAAAGTTGCA 2202  
Qy 640 ValIleGluAspAspSerAspGluGlyIyAspAspGluTyrCysAspGlyAspGluAsp 659  
Db 2203 GTCTCTC--AATGACAGCGCAAGAGTGTGATGACGAGGCCACGAGTTGCAATGACCAACTT 2259  
Qy 660 GluAspAspLeuIySGlySGlyProLeuIySGlyLeuAspGluThrAspArgPheLeuMetThrLeu 679  
Db 2260 AAGGGCGATGTAAAGAAATCTTTGAAACTGACGAAACGAGTGAATTTCTCATGAATTTG 2319  
Qy 680 PhePheGluAsn 683  
Db 2320 TTCTTTGATTAAC 2331

RESULT 5  
BQ233683 1013 bp mRNA linear EST 02-MAY-2002  
LOCUS BQ233683  
DEFINITION AGENCOURT 7258428 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5786526  
5', mRNA sequence.  
ACCESSION BQ233683  
VERSION BQ233683.1 GI:20415083

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.  
REFERENCE 1 (bases 1 to 1013)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LHAM12875 row: 9 column: 07  
High quality sequence stop: 721.  
FEATURES  
source  
1. 1013  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5786526"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb."  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.88e-160 length: 1013  
Score: 1597.00 Matches: 322  
Percent Similarity: 95.6% Conservative: 5  
Best Local Similarity: 94.2% Mismatches: 9  
Query Match: 30.1% Indels: 8  
DB: 3 Gaps: 3  
US-09-515-363C-2 (1-1025) x BQ233683 (1-1013)  
QY 529 LysaAaGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAsp 548  
DB 2 AAAAACCATAATACAGAGCCATGCAGAAGTTGCCATTGCAGATGCACCAAGAGAAGAT 61  
QY 549 ProPheLysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSer 568  
DB 62 CCATTAAAGAGAACTTCTAGAAATATATGACAAAGATTCAAACTTATTGTCAAATGAGT 121  
QY 569 PrometSerAspPheGlyThrGlnProTyrGlnGlnIleGlnIleGlnMetGluLysLys 588  
DB 122 CCAATGTCAGATTCTTGAACTCAACCTATGAAACAATGGCCATTCAAAATGAAAAAAA 181  
QY 589 AlaAlaLysLysGlyAsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsn 608  
DB 182 GCTGCAAAAGAGAAATCGCAAGAAAGTGTGTTGTCAGAAACATTGAGGAAGTACAAT 241  
QY 609 GluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThr 628  
DB 242 GAGGCCCTACAATAATGACACAATTCGAATGATAGATGCGTATACTCATCTTGAAACT 301  
QY 629 PheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspAspSerAspGluGly 648  
DB 302 TTCTATAATGAGAGAAAGATAGAAAGTTTGCAAGTCATAGAAAGATGATAGTGAAGGT 361  
QY 649 GlyAspAspGluTyrCysAspGlyAspGluAspGluAspAspLeuLysProLeuLys 668  
DB 362 GGTGATGATGAGTATTGTGATGGTGAATGAAGATGAGGATGATTTAAAGAAACCTTTGAAA 421

QY 669 LeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLys 688  
DB 422 CTGATGAAAACAGATAGATTCTCATGACTTATTTTGGAAAAACAATAAATGTGAAA 481  
QY 689 ArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIle 708  
DB 482 AGGCTGGCTGAAAAACCCAGAAATATGAAAAGCTGACCAAAATTAAGAAATACCAT 541  
QY 709 MetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArg 728  
DB 542 ATGAGCAATATACTAGACTGAGGAATCAGACGAGGAATATCTTACAAAAACACGA 601  
QY 729 GlnSerAlaTyrAlaLeuSerGlnTyrIleThrGluAsnGluLysPheAlaGluValGly 748  
DB 602 CAGAGTCATATGCGCTTCCCAAGTGAATTAATAAATGAATAATTGCTGAAGTAGA 661  
QY 749 ValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGln 768  
DB 662 GTCAAAAGCCACCATCTGATTGAGCTGACACAGCAGTGAAGTTCAAAACCATGACACAG 721  
QY 769 AsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAla 788  
DB 722 AATGAACAAAAGAAAGTCAATTAGTAATTTCCGACTGAAAAAATAAATCTGCTTATCGCT 781  
QY 789 ThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGly 808  
DB 782 ACCAGAGTGGCAGAGAAAGTCTGATATTAAAGAATGAACATGTTATCCGTTATGCT 841  
QY 809 LeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSer 828  
DB 842 CTCGTAC-AATGAATAGCCATGCTCCAGGCCCGCTGTCGAGCCAGAGCTGATGAGAGC 900  
QY 829 ThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHisGluThr-----Val 846  
DB 901 ACCTAGCTCTGTTGCTTCACGTGGTTCA-----GAGTATCGAAGCTGAGACATT 951  
QY 847 AsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsnMetLysPro 866  
DB 952 AATGATTTCCGAGA-AAAGATGATGATTAAG---CTATACATGGGTTCAAAATGAAACCG 1007  
QY 867 GluGlu 868  
DB 1008 GAGGAG 1013  
RESULT 6  
BQ960157 870 bp mRNA linear EST 21-AUG-2002  
LOCUS  
DEFINITION AGENCOURT\_8949486 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6472264  
5', mRNA sequence.  
ACCESSION BQ960157  
VERSION BQ960157.1 GI:22375635  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.  
REFERENCE 1 (bases 1 to 870)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LHAM14005 row: 0 column: 17  
High quality sequence stop: 738.  
FEATURES  
Location/Qualifiers

source		1..870			
		/organism="Homo sapiens"			
		/mol_type="mRNA"			
		/db_xref="taxon:9606"			
		/clone="IMAGE:6472264"			
		/tissue_type="leiomyosarcoma"			
		/lab_host="DH10B (phage-resistant)"			
		/clone_lib="NH_MGC_71"			
		/note="Organ: uterus; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."			
ORIGIN					
Alignment Scores:					
Pred. No.:		1.7e-136	Length:	870	
Score:		1376.00	Matches:	279	
Percent Similarity:		97.9%	Conservative:	1	
Best Local Similarity:		97.6%	Mismatches:	2	
Query Match:		25.9%	Indels:	5	
DB:		5	Gaps:	0	
US-09-515-363C-2 (1-1025) x BQ960157 (1-870)					
Qy	592	LyegLyAsnArgLySGluArgValCySAAlaGluHisLeuArgLySTyrAsnGluAlaLeu	611		
Db	3	GAAGGAATCGCAAGAACGCTGTTGTGCAGAAAT-TTGAGGAAGTACAAATGAGGCCCTA	61		
Qy	612	GlnIleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsn	631		
Db	62	CAAAATTAATGACACAAATTCGAATGATGATCGGTATCTCATCTTGAACCTTCTATAT	121		
Qy	632	GluGluLyAspLySLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAsp	651		
Db	122	GAAGAGAAAGATAGAGAGTTTGCAGTCATAGAGAGATGATGAGGGTGTGATGAT	181		
Qy	652	GluTyrCyAspGlyAspGluAspGluAspAspLeuLySLysProLeuLySLysLeuAspGlu	671		
Db	182	GAGTATTGTGATGTGATGAAGATGAGGATGATTTAAAGAAACCTTGAACCTGATGAA	241		
Qy	672	ThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLySMetLeuLyArgLeuAla	691		
Db	242	ACAGATAGATTCTCATGACTTATTTTGAAGAACATAAATGTTGAAGAGCTGGCT	301		
Qy	692	GluAsnProGluTyrGluAsnGluLySLysLeuThrLySLysLeuArgAsnThrIleMetGluGln	711		
Db	302	GAAGAACCCAGATATGAAATGAAAAAGCTGACCAATTAAGAATAACCATATGAGACAA	361		
Qy	712	TyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLySThrArgGlnSerAla	731		
Db	362	TATACTAGGACTGAGGAATCAGCAGGAAATATCTTTACAAAAACAGACAGAGTGCA	421		
Qy	732	TyrAlaLeuSerGlnTyrIleThrGluAsnGluLySLysPheAlaGluValGlyValLySAla	751		
Db	422	TATGCCCTTCCAGTGATTTACTGAAATGAAAAATTTGCTGAAGTAGAGTCAAAAGCC	481		
Qy	752	HisHisLeuIleGlyAlaGlyHisSerSerGluPheLySPrometThrGlnAsnGluGln	771		
Db	482	CACCATCTGATTGGAGCTGACACAGCAGTGAAGTCAAAACCATGACACAGATGAACAA	541		
Qy	772	LySGluValIleSerLySPheArgThrGlyLySLysIleAsnLeuLeuIleAlaThrThrVal	791		
Db	542	AAAGAGTCATTAGTAATTCGCACTGGAATAAATCTGCTTATCGCTACCAAGTG	601		
Qy	792	AlaGluGluGlyLeuAspIleLySLysGluCyAsnIleValIleArgTyrGlyLeuValThr	811		
Db	602	GCAGAGAGAGGTCTGGATATTAAAGAAATGTAACATTGTTATCCGTTATGCTCTCAC	661		
Qy	812	AsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrVal	831		
Db	662	AATGAATAGCATGTGTCAGGCCCGTGTGAGCCAGAGCTGATGAGAGACCTACGTC	721		
Qy	832	LeuValAlaHisSer-GlySerGlyValIleGluHisGluThrValAsnAspPheArgGlu	851		

Db	722	CTGGTTGCTCACAGTGGGTTTCAGAGTTATCGAACGTGAGACAGTTAATGATTTCCGAGA	781
Qy	851	uLySmetMetTyrLyAlaIleHisCyValGlnAsnMet-LysProGluGluTyr--Al	870
Db	782	GAAGATGATGTATAAGCTATACATTGTGTTCAAAATATGAAACAGAGAGATATGGC	841
Qy	870	ahisLySile	873
Db	842	TCATAAGATT	851
RESULT 7			
CR857114			
LOCUS	CR857114	2530 bp	mRNA linear HTC 12-NOV-2004
DEFINITION	Pongo pygmaeus mRNA; cDNA DKFZp46900632 (from clone DKFZp46900632).		
ACCESSION	CR857114		
VERSION	CR857114.1	GI:55725101	
KEYWORDS	HTC.		
SOURCE	Pongo pygmaeus (orangutan)		
ORGANISM	Pongo pygmaeus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pongo.		
REFERENCE	1 (bases 1 to 2530)		
AUTHORS	Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.		
CONSRMT	The German cDNA Consortium		
TITLE	Direct Submission		
JOURNAL	Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY		
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp46900632) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp46900632 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.		
FEATURES			
source			
	1. .2530		
	location/Qualifiers		
	/organism="Pongo pygmaeus"		
	/mol_type="mRNA"		
	/db_xref="taxon:9600"		
	/clone="DKFZp46900632"		
	/tissue_type="kidney"		
	/clone_lib="469 (synonym: pkid1). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"		
	/dev_stage="adult"		
	/note="hypothetical protein (Homo sapiens)"		
gene	1. .2530		
	/gene="DKFZp46900632"		
CDS	121. .2157		
	/gene="DKFZp46900632"		
	/codon_start=1		
	/product="hypothetical protein"		
	/protein_id="CAH89418.1"		
	/db_xref="GI:55725102"		
	/translation="MELRPYQWEVIMPALGKNIILPTGAGKTRAAYVAKRHLLETVDGKVVLVNVAHLVTQGEFRRMLDGRWTMTTLLSGDMGPRAGFGLARCHDLITC TABLQMALTSPEEBEHLLETAFSLIVDECHHTKOTVYNNVIMSQYLBKLQRAOPL PVLGLTASPGTGASKLNGAINHVLQCANLDTWCIMS PONYCPQLQESHQQPCQY NLCHRRSQDPGDLKLKMDQIHDLMEBELSRNFGTQMYEQVYKLSAALAGLOE QRVYALHLRRYNDALLIHDTRAVADALALQDFYHREHVTKTQILCAKRLLALFDDH KNVLAHLATHGPNPKLEMLEKILQRFRRSSDSPRGIIFTTRQSAHSLMLMQQPG LQTVDIRAQLIGAGNSSQSTHMTQEQEVIRKFRDGTNLNLTAVTSVAEGLDIPHC NVVRYGLTINEISMVQARGRADQSVYSFVATGSERELKRELINAELETLMEOAVA AVQKMDQAEYQAKIRDLQOALTKRAQAQARENQRQFVEHYQLCINCMVAVGHG SDRKVEGTHVNVNPSFSIYYNVSDDPVINKVFKDWKPGGYISCRNGBVWGLQMI YKSVKLPALKVRSMLLETPOGRIOAKKMSRVFPSPDFLQHCQNLSDLSLD"		
ORIGIN			
Alignment Scores:			

Pred. No. : 1.1e-134 Length: 2530  
Score: 1366.00 Matches: 294  
Percent Similarity: 60.3% Conservative: 138  
Best Local Similarity: 41.0% Mismatches: 233  
Query Match: 25.7% Indels: 52  
DB: 4 Gaps: 12

US-09-515-363C-2 (1-1025) x CR857114 (1-2530)

Qy 306 LeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAsnIle 325  
Db 121 ATGAGCTGCGGCCCTACAGTGGAGGTGATCATGCTGCTGAGGCGCAAGATATC 180  
Qy 326 IleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIleAlaLysAsp 345  
Db 181 ATCATCTGGCTGCCACCGGTCGGGAAGACCCGGCGGCTGCTTATGTGCCCAAGAG 240  
Qy 346 HisLeuAspLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLys 365  
Db 241 CACCTAGAG-----ACTGTGGATGAGGCAAGTGTGTATGTGTCACACAG 288  
Qy 366 ValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeuLysTrpTyr 385  
Db 289 GTGCACCTGGTGACCCAG---CATGGTAAGAGTTCAAGCGCATGTGATGACGCTGG 345  
Qy 386 ArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPheProGluValLys 405  
Db 346 ACCATGACAACCTGAGTGGGACATGGACCACTGTGCTTGGCCACCTGGCCCG 405  
Qy 406 SerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGlu 425  
Db 406 TGCCATGACCTGCTCATCTGCACGGCAGAGCTTTCGACATGGCACCTGACCCCGAG 465  
Qy 426 AsnGlyLysAspAlaGlyValGlnLeuSerAspPheSerLeuIleIleAspGluCys 445  
Db 466 -----GAGGAGGAGCACTTGAGCTCACTGCCCTTCTCCCTGATCGTGTGATGAGTGC 519  
Qy 446 HisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGlnLys 465  
Db 520 CACCACACGCACAAAGACACCGCTCTACAACGTATCATGACCCAGTACCTAGAACTTAA 579  
Qy 466 LeuLysAsnAsnArgLeuLysGluAsnLysProValIleProLeuProGlnIleLeu 485  
Db 580 CTCACAGGGCACAG-----CCGCTACCCACGGTGTG 612  
Qy 486 GlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAlaLysAlaGluHis 505  
Db 613 GGTCTCACAGCTCCCCAGGCACTGGCGGGCTCCAACTCAATGGGGCATCAACAC 672  
Qy 506 IleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThrValLysGluAsnLeu 525  
Db 673 GTCTGTCAAGCTCTGTGCCAATTGGACACGTGTGCATGTCAACCCCAAGAACTACTGC 732  
Qy 526 AspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThr 545  
Db 733 CCCCACTGCAGAGACACAGCAAGCTTGCAAACAGTACAACCTCTGTCAAGGCGC 792  
Qy 546 ArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyrCys 565  
Db 793 AGCCAGATCCGTTTGAGACTTGTGAAGCTCAAGAGCTCAATCCAAATCCAGCACCACTG 852  
Qy 566 GlnMetSerPrometSer---AspPheGlyThrGlnProTyrGluGlnTPrAlaIleGln 584  
Db 853 GAGATGCTGTAGTTGAGCCGGAACCTTGGGACGCAAGATGTATGAGCAGAGTGGTGAAG 912  
Qy 585 MetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCysAlaGluHisLeu 604  
Db 913 TTGAGTGAAGCTTGGCTTGGCCGGGCTCCAGAGCAACGGGTGTATGCGCTTCACTG 972  
Qy 605 ArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyrThr 624  
Db 973 AGGCGCTACAATGACGCGCTGCTCATCATGACACCGTCCGCGCGGTGATGCTTGGCT 1032

Qy 625 HisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspAsp 644  
Db 1033 GCGCTGCAGAGATTCTATTCACAGGAGCATGTCACTAAACCACAGATCTG----- 1083  
Qy 645 SerAspGluGlyLysAspAspGluTyrCysAspGlyAspGluAspGluAspLeuLys 664  
Db 1084 -----TGT----- 1086  
Qy 665 LysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsn 684  
Db 1087 -----GCCAAGCGCGGCTGCTGGCCCTGTTCGATGACACCAAG 1125  
Qy 685 LysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeu 704  
Db 1126 AATGTCTGGCCCACTGGCACTCATGGCCCA---GAGATCCAAACTGGAAGATGCTG 1182  
Qy 705 ArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePhe 724  
Db 1183 GAAAGATCCTGCAAGGCAAGTTC--AGAGCTCTGACAGCCCTCGGGGTATCATCTTC 1239  
Qy 725 ThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTPrIleThrGluAsnGluLysPhe 744  
Db 1240 ACCCGACCCCGCAGAGGACACTCCCTCTGCTGTGCTCCAGCAGACCGAGCCTG 1299  
Qy 745 AlaGluValGlyValLysAlaHisHisLysLeuIleGlyAlaGluHisSerSerGluPheLys 764  
Db 1300 CAGACCGTGACATTGGGCCCCAGCTACTGATGGGGCTGGGAACAGACGACAGACACC 1359  
Qy 765 PrometThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIleAsn 784  
Db 1360 CACATGACCCAGAGGACCAAGAGAGTGTCCGGAAGTTCGAGATGGAAACCTGAAC 1419  
Qy 785 LeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleVal 804  
Db 1420 CTTCTGTGTGCCACAGTGTGGCGGAGGAGGGCTGACATCCCACTTGCATGTGTG 1479  
Qy 805 IleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArg 824  
Db 1480 GTGGCTTATGGGCTCTGCACCAATGAATCTCATGTGTCCAGGCGAGGGCGGTGCCCG 1539  
Qy 825 AlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHisGlu 844  
Db 1540 GCGGATCAGAGTGTACTCGTTTGTAGCACTGAGGGTAGCCGGAGCTGAACGGGAG 1599  
Qy 845 ThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsnMet 864  
Db 1600 CTGATCAACGAGGCGCTGAGACGCTGATGGACAGGCACTGCTGTGCAGAAATG 1659  
Qy 865 LysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGluLys 884  
Db 1660 GACCAGGCTGAGTACAGGCCCAAGATCCGGGATCTGCAGCAGGCGAGCCCTGACCAAGCGG 1719  
Qy 885 LysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIleThr 904  
Db 1720 GCGGCCACGCGACGCCGCGGAGAAACCAAGCGGCGAGTTCACAGTGAGCAGCTG 1779  
Qy 905 PheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIleGlu 924  
Db 1780 CTACTCTGCATCAACTGCATGTGGCGCTGGGCCACGCGCAGCACTGCGGAAGTGAG 1839  
Qy 925 LysMetHisHisValAsnMetThrProGluPheLysGluLeuTyr---IleValArgGlu 943  
Db 1840 GGCACCCACCATGTCAACGTGAACCCCAAGCTTCCGATCTACTATATGTCTCCAGGAT 1899  
Qy 944 AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLys 963  
Db 1900 CCTGTGTCATCAACAAAGTCTTCAAGGACTGGAAGCCTGGGGGTGTCACTGACG 1959  
Qy 964 ---CysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeu 982  
Db 1960 AACTGTGGGAGGTCTGGGGTCTGCAGATGATCTCAAGTCAAGTGAAGCTGCGCGCTC 2019  
Qy 983 LysIleArgAsnPheValValValPheLysAsnAsnSerThrLysLysGlnTyrLysLys 1002

||||:||||: :||| :|||  
Db 2020 AAAGTCGCAGC-----ATGCTGTGAGAGCCCTCAGGGGCGGATCCAGGCCAAAAAG 2073  
Qy 1003 TrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCys 1019  
||| :||| :||| :||| :|||  
Db 2074 TGGTCCCGCGGTCCCTTCCTCGCTGACTTTGACTTCTCGACAGATTGT 2124  
RESULT 8  
BU189982  
LOCUS BU189982 918 bp mRNA linear EST 04-SEP-2002  
DEFINITION AGENCOURT\_7950921 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6173848  
5' mRNA sequence.  
ACCESSION BU189982  
VERSION BU189982.1 GI:22703966  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 918)  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Place: LLAM13546 row: a column: 17  
High quality sequence stop: 556.  
location/Qualifiers  
1. 918  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6173848"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. primer: Oligo dT.  
Average insert size 2.1 kb."  
ORIGIN  
Alignment Scores: 3.64e-127 Length: 918  
Pred. No.: 1289.50 Matches: 269  
Score: 95.1% Conservative: 5  
Percent Similarity: 93.4% Mismatches: 6  
Best Local Similarity: 24.3% Indels: 8  
Query Match: 5 Gaps: 1  
US-09-515-363C-2 (1-1025) x BU189982 (1-918)  
Qy 360 lIeValleuValAsnLyseValleuLeuValGluGlnLeuPheArgLyseGluPheGlnPro 379  
|||||  
Db 2 ATAGTTCTTGTCAATAAGGTACTGCTAGTGTGAACAGCTCTTCGCAAGAGGTTCACACCA 61  
Qy 380 PheLeuLySLySTrPTyrArgValIlleGlyLeuSerGlyAspThrGlnLeuLySIlleSer 399  
|||||  
Db 62 TTTTGAAGAATGATCGTATCGTTATTGATTAGTGGTGATACCCCAACTGAAAAATATCA 121  
Qy 400 PheProGluValValLySerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsn 419  
|||||  
Db 122 TTTCAGAGAAGTGTCAAGTCTCTGTGATATTATTATCATGACAGCTCAAAATCCTTGAAAAAC 181  
Qy 420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439  
|||||

Db 182 TCCCTCTTAACCTGGAAAAATGAGAGAATGCTGTGTTCAATTGTCAAGACTTTTCCCTC 241  
Qy 440 lIeIleIleAspGluCysHisIleThrAsnLySGluAlaValTyrAsnAsnIleMetArg 459  
|||||  
Db 242 ATTATCATGTGATGATGTCATCATCACACCAACAAGAAGCAGTGTTAATAACATCATGAGG 301  
Qy 460 HisTyrLeuMetGlnLySLeuLySAsnAsnArgLeuLySlySGluAsnLySProValIle 479  
|||||  
Db 302 CATATTGTGATGCAGAAAGTTGAAAAACATAGACTCAAGAAAGAAAAACAACAGTGATT 361  
Qy 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLySGln 499  
Db 362 CCCCTTCTCAGATACTGGGACTAACAGCTTCACTGTGTTGAGGGGCCACGAAGCAA 421  
Qy 500 AlAlYSAlaGluGluHisIleLeuLySLeuCysAlaAsnLeuAspAlaPheThrIleLyS 519  
Db 422 GCCAAAGCTGAAGACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAA 481  
Qy 520 ThrValLySGluAsnLeuAspGlnLeuLySAsnGlnIlleGlnGluProCysLySlySPh 539  
Db 482 ACTGTTAAGAAACCTTGATCACTGAATAAACCAATACAGAGCATGCAAGAAGTTT 541  
Qy 540 AlaIleAlaAspAlaThrArgGluAspProPheLySGluLySLeuLeuGluIleMetThr 559  
Db 542 GCCATTGCAGATGCACCAAGAGATCCATTTAAGAGAACTTCTAGAAATATGACA 601  
Qy 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579  
Db 602 AGGATTCAAACTTATTGTCAATGAGTCCAAATGTCAAGATTGTGAACTCAACCTATGAA 661  
Qy 580 GlnTrpAla--IlleGlnMetGluLySlySAlaAlaLySlySGlyAsnArgLySGluArg 598  
Db 662 ACATGGGCCCATTTCCAATGAAAAAAAGCCTGCAAAAAGAAATCGCAAGAACGT 721  
Qy 599 Val-CysAlaGluHisLeuArgLySTyrAsnGlu-AlaLeuGlnIle-AsnAspThrIle 617  
Db 722 GTTTTGTGCAAAACATTTGAGGAAGTACATGAAGGCCCTACAAATTAAATGACACAAATT 781  
Qy 618 ArgMet-IleAspAlaTyrThrHisLeu-GluThrPheTyrAsn-GluGluLySAsp-Ly 636  
Db 782 TCAATGATAAATGCGTATACTCATCTTTGAACCTTCTATATATGAGAAAGAGATTAG 841  
Qy 636 slySPheAla 639  
Db 842 AAAGTTTGCC 851  
RESULT 9  
BU902097 1174 bp mRNA linear EST 17-OCT-2002  
LOCUS BU902097  
DEFINITION AGENCOURT\_10127740 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6502757  
5' mRNA sequence.  
ACCESSION BU902097  
VERSION BU902097.1 GI:24084010  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 1174)  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Place: LLAM14059 row: f column: 06

High quality sequence stop: 808.  
Location/Qualifiers  
1..1174  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6502757"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1ib="NIH MGC 71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb. "  
  
ORIGIN  
  
Alignment Scores:  
Pred. No.: 1.67e-123 Length: 1174  
Score: 1257.00 Matches: 271  
Percent Similarity: 87.8% Conservative: 10  
Best Local Similarity: 84.7% Mismatches: 19  
Query Match: 23.7% Indels: 20  
DB: 5 Gaps: 1  
  
US-09-515-363C-2 (1-1025) x BU902097 (1-1174)  
  
Qy 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20  
Db 223 ATGTCGAATGGGTATTCACAGACGAGAATTCGCTATCTCATCTCGTGTTCAGGGCC 282  
  
Qy 21 ArgValIysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40  
Db 283 AGGGTGAATAATGTACATCCAGGTGAGCCTGTGTGACGACTACCTGACCTTCTGCCTGCA 342  
  
Qy 41 GluValIysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60  
Db 343 GAGGTGAAGGACGAGATTCCAGACGACAGTCCGCACTCCGGGAACATGCAGGCA GTTGA 402  
  
Qy 61 LeuLeuLeuSerThrLeuGluIlySGlyValTPrHisLeuGlyTPrThrArgGluPheVal 80  
Db 403 CTGCTGCTGAGCACCTTGAGAGGAGTCTGGCACCTTGCTTGACTCGGGAATTCTGTG 462  
  
Qy 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100  
Db 463 GAGGCCCTCCGAGAACCGGACCCCTCTGGCCGCCGCTACATGAACCTGAGCTCAGC 522  
  
Qy 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120  
Db 523 GACTTGCCCTCTCCATCGTTGAGAACGCTCATGATGATATCTCCAACCTGCTGAACCTC 582  
  
Qy 121 LeuGlnProThrLeuValAspIlySleuLeuValArgAspValLeuAspIlyCysMetGlu 140  
Db 583 CTTACAGCCCACTCTGCTGACAAAGCTTCTAGTTAGACGCTCTTGATAGTCATGGAG 642  
  
Qy 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160  
Db 643 GAGGAACCTGTTGACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAAACAATGGAAT 702  
  
Qy 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTPrPheSerAla 180  
Db 703 GAATCAGGTGTAGAGAGCTAATAAAAGATTGTGCAGAAAGAAAAGTGTCTCTGCA 762  
  
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200  
Db 763 TTTCTGAATGTTCTTCTGCAACAGGAAACAATGAATGTTGCCAAGAGTTAACAGGCTCT 822  
  
Qy 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220  
Db 823 GATTGCTCAGAAAGCAATGCAGAGATTGAGAATTATCACAAGTTGATGCTCTCAAGTG 882  
  
Qy 221 -GluGluGlnLeuLeuSerThrThr-ValGlnProAsnLeuGlu-LysGluValTPrGly 239  
Db 883 GGAAGAACCACTTCTTTCAACCAACAGGTTCAGCCAAATCTGAGAAAAAGGAGGTCTGGGGC 942

Qy 240 MetGlu-AsnAsnSerSer-GluSerSerPheAlaAspSerSerValValSerGluSera 259  
Db 943 ATGGAATAAATACTCATCAAAATAATCATCTTTGCAGATCCTTCTGAAGTTCCAGAACCA 1002  
  
Qy 259 sPrThrSerLeuAla-GluGlySerValSerCysLeu-AspGluSer-LeuGlyHisAsn 278  
Db 1003 AACCAAGTTTGGCAAAAGAAATGCAACTGCTTAAATGAAGTCTTGGACATACAA 1062  
  
Qy 278 exAsnMetGlySerAsp-----SerglyT 286  
Db 1063 ACAACATGGGCAGGAGATTCCGGCCCCGGGGCAAGATTTCAAATTAATAAAAGGGG 1122  
  
Qy 286 hMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300  
Db 1123 CCACCAAGAAATTCGCCGGAACAAAAAATTCACGCTTGGGCC 1166  
  
RESULT 10  
BM467983 1115 bp mRNA linear EST 05-FEB-2002  
LOCUS  
DEFINITION AGENCOURT\_6437921 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5532884  
5', mRNA sequence.  
ACCESSION BM467983  
VERSION BM467983.1 GI:18517025  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1115)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL2216 row: n column: 21  
High quality sequence stop: 690.  
  
FEATURES  
source  
location/Qualifiers  
1..1115  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5532884"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1ib="NIH MGC 71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb. "  
  
ORIGIN  
  
Alignment Scores:  
Pred. No.: 2.39e-119 Length: 1115  
Score: 1218.00 Matches: 228  
Percent Similarity: 99.1% Conservative: 0  
Best Local Similarity: 99.1% Mismatches: 2  
Query Match: 22.9% Indels: 0  
DB: 3 Gaps: 0  
  
US-09-515-363C-2 (1-1025) x BM467983 (1-1115)  
  
Qy 796 LeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAla 815  
Db 1 TTGGATATTAAAGAAATGAACATGTTATCCGTTATGCTCGTCAACCAATGAATAGCC 60  
  
Qy 816 MetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHis 835

Db	61	ATGCTCCAGGCCCGTGTGAGCCAGAGCTGATGAGACACCTACGTCTCGGTGCTCAC	120
Qy	836	SerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluIuysMetMetTyr	855
Db	121	AGTGGTTCAGAGAGTATTCGACGTCGAGACAGTAAATGATTTCCGAGAGAGATGATGAT	180
Qy	856	LysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisIleLysIleLeuGlu	875
Db	181	AAAGCTATACATTTGTGTTCAAAATATGAAACCAGAGAGATATGCTCATAGATTTTGAA	240
Qy	876	LeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHis	895
Db	241	TTACAGATGCMAAGTATATGAGAAAAGAAATGAAACCAAGAGAAATATTGCCAAGCAT	300
Qy	896	TyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValIleuAlaCys	915
Db	301	TACAAGAATAACCCATCACTAATAACTTCTCTTGCAAAAACCTGCAGTGTGCTAGCCTGT	360
Qy	916	SerGlyGluAspIleHisValIleGluLysMetHisHisValAsnMetThrProGluPhe	935
Db	361	TCTGGGGAAGATATCCCATGTAATTGAGAAATGCATCAGTCATATGACCCAGAATTTC	420
Qy	936	LysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGln	955
Db	421	AAAGAACTTTACATTTGTAGAGAAAAACAACACACTGCAGAAAGAGTGCCGACTATCAA	480
Qy	956	IleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrrGlyThrMetMetValHisLys	975
Db	481	ATTAATGCTGAATCATCTGCAGAAATGTGGCCAGGCTTGGGGAACAATGATGGTGACAAA	540
Qy	976	GlyLeuAspLeuProCysLeuLysIleArgAsnPheValValPheLysAsnAsnSer	995
Db	541	GGCTTAGATTGGCTTGTCTCAAAATAGGAATTTGTAGTGGTTTCAAAAATAATTCA	600
Qy	996	ThrLysLysGlnTyrLysLysTrrValGluLeuProIleThrPheProAsnLeuAspTyr	1015
Db	601	ACAAGAACAATACAAAGAGGGTAGAATTACCTATCACATTTCCCAATCTTGACTAT	660
Qy	1016	SerGluCysCysLeuPheSerAspGluAsp	1025
Db	661	TCAGAAATGCTGTTATTAGTATGATGAGGAT	690
RESULT 11			
BF337464			
LOCUS	BF337464	1035 bp	mRNA linear EST 22-NOV-2000
DEFINITION	602035195F1 NCI CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4183126		
5', mRNA sequence.			
ACCESSION	BF337464		
VERSION	BF337464.1	GI:11283715	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Homidae; Homo.			
REFERENCE	1 (bases 1 to 1035)		
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a> Tissue Procurement: David N. Louis, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Place: LAM9498 row: n column: 23 High quality sequence start: 5 High quality sequence stop: 695.		
FEATURES			
Location/Qualifiers			

source

1. .1035

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4183126"

/issue\_type="g1loblastoma with EGFR amplification"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NCI CGAP\_Brn64"

/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.57 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

3.65e-118

1206.50

84.0%

79.5%

22.7%

2

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

1035

264

15

44

12

5

US-09-515-363C-2 (1-1025) x BF337464 (1-1035)

Qy

295

AsnValAlaAlaArgAlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGlu

314

Db

6

AATGTGGCAGCAGACATCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAA

65

Qy

315

ValAlaGlnProAlaLeuGluGlyLysAsnIleIleCysLeuProThrGlySerGly

334

Db

66

GTTGCCAGCCAGCCTTGGAAGGGAAGTAATCATCATCTGCTCCCTTACAGGAGTGA

125

Qy

335

LysThrArgValAlaValTyrIleAlaLysAspHisLeuAspLysLysAlaSer

354

Db

126

AAAACCAAGTGGCTGTTTACATTGCCAAGATCACTTAGACAAGAAAAAGCATCT

185

Qy

355

GluProGlyLysValIleValLeuValAsnLysValLeuLeuValGluGlnLeuPheArg

374

Db

186

GAGCCTGGAAGAAGTTATAGTCTTGTCATTAAGTACTGCTAGTGAACAGCTCTCCGC

245

Qy

375

LysGluPheGlnProPheLeuLysLysTrrValArgValIleGlyLeuSerGlyAspThr

394

Db

246

AAGGAGTTCCAACCATTTTGAAGAAATGTAATCGTGTATTGATTAAGTGTGATACC

305

Qy

395

GlnLeuLysIleSerPheProGluValValLysSerCysAspIleIleIleSerThrAla

414

Db

306

CAACTGAATATCAATTTCCAGAAAGTTCAGAGTCTGTGATATTATTATCAGTACAGCT

365

Qy

415

GlnIleLeuGluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeu

434

Db

366

CAATCCTTGAAACCTCCCTTTAAACTTGCAAAATGAGAAAGATGCTGTTCATTTG

425

Qy

435

SerAspPheSerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyr

454

Db

426

TCAGACTTTTCCCTCATTTATCATTTGATGATGATGCATCACACCAACAAGACAGTAT

485

Qy

455

AsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGlu

474

Db

486

AATAACATCATGAGGCAATTATTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGA

545

Qy

475

AsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGly

494

Db

546

AACAAACAGTATTCCTTCTCAGATAGTGGACTAACAGCTTCACTGTGTGCGA

605

Qy

495

GlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAsp

514

Db

606

GGGGCCACGAAGCAAGCCAAAGCTGAAGACACATTTTAAACTATGTGCCAATCT-GAT

664

Qy

515

AlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGlu

534

Db

665

GCAATTACTATTAACCTGTGTAGAAAACCTTGATCACTGGGAAAAACAATACAGAGG

724

Qy

535

ProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeu

554

Db 725 CCAATCGAGAAGTTGGCGCT-GCGGATGCGAC-CGAGAAGATCCCTTTAAGGGAAC--- 779  
Qy 555 LeuGluIleMetThrArgIleGlnThrTyrcysGlnMetSerPro-MetSerAspPheG1 574  
Db 780 ---TCTGAATAATAGGCAGGGTTCCACTTTGGTGCACT---AGTCCAGGTCCGA---TTGGG 830  
Qy 574 YThrGlnProTyrcysGlnIleTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAs 594  
Db 831 TAGCCACCCTATTATAA-----GGGGCCTCAACGGGGGAAAGGGGCTAAGAAACACAA 884  
Qy 594 nArgLysGluArgValCysAlaGluHisLeuArgLysTyraSngLysAlaLeuGlnIleAs 614  
Db 885 ACCGTTGGCGACCTTGGGGCGAACAAGCGCACAAAGACACAGAAAGCGCACAA 944  
Qy 614 nAspThrIleArgMetIleAspAlaTyrrHis 625  
Db 945 CGAACACAGCAA-----GACAGAACGACACAC 972

RESULT 12

LOCUS DR422160 690 bp mRNA linear EST 29-JUN-2005  
DEFINITION nav08g05.y1 Human pterygium. Unnormalized (nav) Homo sapiens cDNA  
clone nav08g05 5', mRNA sequence.

ACCESSION DR422160  
VERSION DR422160.1 GI:68324176  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE 1 (bases 1 to 690)  
AUTHORS Wistow,G., Jaworski,C., Aryankalayil-John,M., Rowsey,J.J., Cox,C.,  
Reid,T., Dushku,N. and Carper, D.  
TITLE NEIBank analysis of Human pterygium  
JOURNAL Unpublished (2005)  
COMMENT Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 08 row: g column: 05  
Seq primer: Universal M13 Reverse.

FEATURES

source

1..690  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="nav08g05"  
/tissue\_type="Pterygium"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="Human pterygium. Unnormalized (nav)"  
/note="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted  
from 9 pooled human pterygia. A directionally cloned cDNA  
library in the pCMVSPORT6 vector (Invitrogen) was  
constructed at Bioserve Biotechnology (Laurel MD)  
essentially following the protocols of the SuperScript  
Plasmid System, full details of which are contained in the  
manufacturer's instruction manual  
(http://www.lifetech.com/). First strand synthesis was  
carried out using a Not I primer-adapter  
[5'-pGACTAGTCTAGATCGGAGCGCGCC(T)15-3']. cDNA was  
cloned in Not I/Sal I sites. EST analysis was performed at  
the NIH Intramural Sequencing Center (NISC). Analyzed data  
available through http://neibank.nei.nih.gov."

ORIGIN

Alignment Scores:  
Pred. No.:

3.06e-116 Length: 690

Score: 1186.00 Matches: 226  
Percent Similarity: 99.1% Conservative: 1  
Best Local Similarity: 98.7% Mismatches: 2  
Query Match: 22.3% Indels: 0  
DB: 8 Gaps: 0

US-09-515-363C-2 (1-1025) x DR422160 (1-690)

Qy 563 ThrTyrcysGlnMetSerProMetSerAspPheGlyThrGlnProTyrcysGlnIleTrpAla 582  
Db 2 ACTTATGTGCTCCCTTAGTCCATGTCAAGATTGGAACTCAACCTATGACATGGGCC 61  
Qy 583 IleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCysAlaGlu 602  
Db 62 ATTCAAATGGAATAAAAGCTGCCAAAGAGAAATCCCAAGAACGTGTTGTGCAGAA 121  
Qy 603 HisLeuArgLysTyraSngLysAlaLeuGlnIleAsnAspThrIleArgMetIleAspAla 622  
Db 122 CATTGAGGAAGTACATGAGGCCCTACAAATTATGACACAAATTCCAAATGATGATGCG 181  
Qy 623 TyrrHisLeuGluThrPheTyraSngLysAlaLysLysLysPheAlaValIleGlu 642  
Db 182 TATACTCATCTTGAACCTTCTATATGAGAGAAAGATAGAGATTGACATAGAA 241  
Qy 643 AspAspSerAspGluGlyLysAspAspGluTyrcysAspGlyAspGluAspGluAspAsp 662  
Db 242 GATGATAGTATGATGAGGGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 301  
Qy 663 LeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGlu 682  
Db 302 TTAAAGAAACCTTGAAACTGATGAAACAGATGATTCTCATGACTTATTTTTGAA 361  
Qy 683 AsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrcysGlnLysLeuThr 702  
Db 362 AACAAATAAATGTTGAAAGGCTGCTGAAACCAGAAATATGAAATGAAAGCTGACC 421  
Qy 703 LysLeuArgAsnThrIleMetGluGlnTyrrHisArgThrGluGluSerAlaArgGlyIle 722  
Db 422 AAATTAAGAAATACCAATAGGACCAATATATCTAGACTGAGGAATCAAGAGAAATA 481  
Qy 723 IlePheThrLysThrArgGlnSerAlaTyrrAlaLeuSerGlnTrpIleThrGluAsnGlu 742  
Db 482 ATCTTACAAAAACAGACAGAGTGCATATGCGCTTCCAGTGCATTAAGAAATGAA 541  
Qy 743 LysPheAlaGluValGlyValLysAlaHisIleLeuIleGlyAlaGlyHisSerSerGlu 762  
Db 542 AAATTGCTGAAGTAGAGTCAAGCCCAACCATCTGATTGGAGCTGCACACAGACTGAG 601  
Qy 763 PheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLys 782  
Db 602 TTCAAACCCATGACACAGATGAACAAAGAGTCAATTAATTTCCACTGGAATA 661  
Qy 783 IleAsnLeuLeuIleAlaThrThrVal 791  
Db 662 ATAAATCTGCTTATCGCTACACAGTG 688

RESULT 13  
LOCUS BF983236 729 bp mRNA linear EST 23-JAN-2001  
DEFINITION 602305873F1 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:4397083 5',  
mRNA sequence.  
ACCESSION BF983236  
VERSION BF983236  
KEYWORDS BF983236.1 GI:12386048  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 729)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10097 row: a column: 20  
High quality sequence stop: 665.  
Location/Qualifiers  
1..729  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4397083"  
/tissue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 88"  
/note="Organ: small\_intestine; Vector: PCMV-SPORT6;  
Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally;  
oligo-dt primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:  
Pred. No.: 4.44e-110 Length: 729  
Score: 1129.00 Matches: 238  
Percent Similarity: 98.4% Conservative: 2  
Best Local Similarity: 97.5% Mismatches: 4  
Query Match: 21.3% Indels: 3  
DB: 2 Gaps: 0

US-09-515-363C-2 (1-1025) x BF983236 (1-729)

QY 140 GUGUGUGUleuLeuThrIleGluAspArgIleAlaAlaGluAsnAngly 159  
|||||  
Db 1 GAGGAGAACTGTTGACAAATTGAAGACAGAAACCGGATTGCTGTCAGAAACAATGA 60

QY 160 AengLuserGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSer 179  
|||||  
Db 61 AATGAATCAGGTGTAAGAGAGCTACTAAAAAGATTGTGCAGAAAGAAATGCTTCT 120

QY 180 AlaPheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGly 199  
|||||  
Db 121 GCATTTCTGAATGTTCTTCGTCAAACAGAAACAATGAACCTTGTCCAAAGATTAAACAGGC 180

QY 200 SerAspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGln 219  
|||||  
Db 181 TCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTATCACAAAGTTGATGCTCTCAA 240

QY 220 ValGluGluGlnLeuLeuSerThrValGlnProAsnLeuGluLysGluValTrpGly 239  
|||||  
Db 241 GTGGAAGAGCAACTTCTTCAACACACAGTTCAGCCAATCTGGAGAGGAGTCTGGGGC 300

QY 240 MetGluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAsp 259  
|||||  
Db 301 ATGGAGAATAACTCATCATCATCTTTTTCAGATTCTTCTGTAGTTTCAGAAATCAGAC 360

QY 260 ThrSerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsn 279  
|||||  
Db 361 ACAAGTTTGGCAGAAAGAGAGTGTCACTGCTTAGATGAAAAGTCTTGACATTAACAGCAAC 420

QY 280 MetGlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArg 299  
|||||  
Db 421 ATGGGAGTGATTCAAGGCACCATGGGAAGTGATTGAGAAAGAAATGTGCAGCAAGA 480

QY 300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319  
|||||  
Db 481 GCATCCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCACGACGAC- 539

QY 320 LeuGluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAla 339  
|||||  
Db 540 TTGGAAGGAGAGAAATATCATCATCTGCTC-CCTACAGGAGCTGAGAAAAACAGAGTGGCT 598

QY 340 ValTyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysVal 359  
|||||  
Db 599 GTTTACATTGCCAAGGATCACTTAGACAG-AAGAAAAAGCATCTGAGCCCTGAAAAAGTT 657

QY 360 IleValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnPro 379  
|||||  
Db 658 ATAGTCTTGTGCAATAGGTAAGTACTGCTAGTTGAACAGCTCTTCGAAAGAGTCCACCA 717

QY 380 PheLeuLysLys 383  
|||||  
Db 718 TTTTGGAGAAAG 729

RESULT 14  
BY720783  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

972 bp mRNA linear EST 17-DEC-2002  
BY720783 RIKEN full-length enriched, adult male cecum Mus musculus  
CDNA clone 9130009C22 5', mRNA sequence.  
BY720783  
BY720783.1 GI:271333900  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 972)  
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,  
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,  
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,  
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,  
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,  
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,  
Chochia,C., Corbani,L.E., Couains,S., Dalla,E., Dragani,T.A.,  
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,  
Gierboldi,M., Glasi,C., Godzik,A., Gough,J., Grimmond,S.,  
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,  
Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,  
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,  
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,  
Numata,K., Okido,T., Pavan,W.J., Perte,G., Pesole,G.,  
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramchandran,S.,  
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,  
Sandelin,A., Schneider,C., Sempile,C.A., Setou,M., Shimada,K.,  
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,  
Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,  
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,  
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,  
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,  
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,  
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,  
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,  
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,  
Rogers,J., Birney,E. and Hayashizaki,Y.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
12466851

TITLE  
JOURNAL  
PUBMED  
COMMENT

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/  
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,  
Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,  
Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,  
Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,

Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source  
1..972  
location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="9130009C22"  
/sex="male"  
/tissue\_type="cecum"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, adult male cecum"  
/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAAAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAATTAATTAATCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from lambda FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"

ORIGIN

Alignment Scores:  
Pred. No.: 3.97e-109 length: 972  
Score: 1122.00 Matches: 215  
Percent Similarity: 92.6% Conservative: 11  
Best Local Similarity: 88.1% Mismatches: 18  
Query Match: 21.1% Indels: 0  
DB: 5 Gaps: 0

US-09-515-363C-2 (1-1025) x BY720783 (1-972)

Qy 782 LysileasnleuleulealathrvalalaglgluglyleuaspilylesgluCys 801  
:::|||||  
Db 1 GAAATAAATCTGCTTATCGCTACGACGGTGGCAGAGAGAGCCCTGATATCAAGAGTGC 60  
Qy 802 Asnilevalileargtyrglyleuvalthrasngluilealamevalglnalaarggly 821  
|||||  
Db 61 AATATTGTATTTCGTTATGGCCTTGTCAAGACGAGATAGCCATGTGCCAGGCCGGGGT 120  
Qy 822 ArgAlaArgAlaAspGluSerThrTyrValleuValAlaHisSerGlySerGlyValIle 841

Db 121 CGAGCCAGAGCTGATGAAGACACGATATGTCTCCGTGCACACGACAGTGGCTCAGAGATTACC 180  
Qy 842 GluHiegluThrvalasnaspPheargGluysmetMetTyrlysalalieHisCysVal 861  
|||  
Db 181 GAACGGAGATTGTTATGATTTCCGAGAGAGATGATGTATTAAGCTATTAACCGTGT 240  
Qy 862 GlnasnMetlysProglugluTyrAlaHislylileleugluLeuglnMetGlnSerIle 881  
Db 241 CAAAACATGAACACAGAGATATGCACATAGATTTTGAAATTGCAGGTGCAAAATATC 300  
Qy 882 MetGluLysLysMetLysThrLysArgAsnilealalyshisTyrlysnasnProSer 901  
Db 301 CTGGAAGAAAGAAATGAAGTCGAAGAGCAATGCAAGCAATACCAATCCATCG 360  
Qy 902 LeuileThrPheleuCysLysasnCysSerValleuAlaCysSerGlygluaspIleHis 921  
Db 361 TTAATAACACTTCTCTGCAAAATGTAGCATGCTGCTGCTCGGAGAAAACATCCAT 420  
Qy 922 ValileglulysMethisHisValasnMetThrProgluPheLysgluLeuTyrIleVal 941  
Db 421 GTCAATTGAGAAGATGCATCATGTCAATATGACACCAAGATTCAAGGAGCTTACATTGTA 480  
Qy 942 ArgGluasnLysAlaleuglnLysLysCysAlaaspTyrGlnileasnlygluileIle 961  
Db 481 AGAGAAAACAAAGCACTGCAGAAAAGAAATTTGCTGATTATCAGACCAATGAGAGATTATC 540  
Qy 962 CysLysCysGlyGlnAlaTrrpGlyThrMetMetValHisLysGlyLeuaspLeuProCys 981  
Db 541 TGCAGTGTGGCCAGGCTTGGGGAACAATGATGTGCACAAAGGTTAGATTGGCTTGT 600  
Qy 982 LeuLysileArgAsnPhelvalValPheLysasnAsnSerThrLysLysGlnTyrLys 1001  
Db 601 CTTAAATAAGAAATTTGTAGTCAATTTCAAAATAACTCACCGAAGAAACAGTACAAG 660  
Qy 1002 LysTrpValGluLeuProIleThrPheProAsnLeuaspTyrSerGlyCysLeuPhe 1021  
Db 661 AAGTGGGTGAATTCCTATTCAGATTTCTGATCTTGACTACTCAGAACTAGCTGTGAT 720  
Qy 1022 SeraspGluasp 1025  
Db 721 AGTGATGAAGAT 732  
RESULT 15  
AK018602 1197 bp mRNA linear HTC 03-APR-2004  
LOCUS  
DEFINITION Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:9130009C22 product:similar to MELANOMA  
DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert sequence.  
ACCESSION AK018602  
VERSION AK018602.1 GI:12858393  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
PUBMED 10349636  
REFERENCE  
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
PUBMED 11042159  
REFERENCE  
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, Y., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,

TITLE	JOURNAL	REFERENCE	AUTHORS
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	RIKEN integrated sequence analysis (RISA) system--384-format	Genome Res. 10 (11), 1757-1771 (2000)	11076861
4			
TITLE	JOURNAL	REFERENCE	AUTHORS
THE RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.	Functional annotation of a full-length mouse cDNA collection	Nature 409, 685-690 (2001)	5
5			
TITLE	JOURNAL	REFERENCE	AUTHORS
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	Nature 420, 563-573 (2002)	6 (bases 1 to 1197)
6			
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.	Direct Submission	Submitted (10-JUL-2000)	Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT			
Please visit our web site (http://genome.gsc.riken.jp/) for further details.			
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGATTCGAGTTAATTAATTAATCCCCCCCCCCC 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATTCGAGTTAATTAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified plasmid KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.			
FEATURES			
source			
1. .1197			
/organism="Mus musculus"			
/mol_type="mRNA"			
/strain="C57BL/6J"			
/db_xref="PANTOM_DB:9130009C22"			
/db_xref="taxon:10090"			
/clone="9130009C22"			
/sex="male"			
/tissue_type="cecum"			
/clone_lib="RIKEN full-length enriched mouse cDNA library"			
/dev_stage="adult"			
<1. .737			
/note="unnamed protein product; putative similar to MEIANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5"			

```

[ Homo sapiens ] (SPTRLQ9HG6, evidence: FASTY, 75.6% ID,
67.6% length, match=2079) "
/codon_start=3
/protein_id="BAB31303.2"
/db_xref="GI:26384184"
/translation="EINLIATVAREGLDIKECNIVIRYGLVTNEIAMVQARGARA
DESTYLVATSSGSGVTEREIVNDRKMYKAINRVQNMKPEEYAHKILEQVSI
KKMKVERSI AKQYNDNPSLITLCKNCSMLVCSGENIHVIEKMHVNTPEFGLYIV
RENKALQKKEPADYQTNGEIICKCGQAWGTWVHKGLDPLCLIRNFVNFKNSPKQ
YKKWVELPIRFPDLDYSEYCLYSDSD"
1180..1185
/note="putative"
1197
/note="putative"
/orig="putative"

ORIGIN

Alignment Scores:
Pred. No.      5.52e-109      Length:      1197
Score:         1122.00      Matches:     215
Percent Similarity: 92.6%      Conservative: 11
Best Local Similarity: 88.1%      Mismatches:  18
Query Match:    21.1%      Indels:      0
DB:             4          Gaps:         0

US-09-515-363C-2 (1-1025) x AK018602 (1-1197)

QY      782  LysLLeAsnLeuLeuLeaLthrThrValAlaGluGluGlyLeuAspIleLeuGluCys
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      3  GAAATAAATCTGCTTAATCGCTACGACGGTGGCAGAGAAAGCCTGATATCAAGATGC
      62

QY      802  AsnIleValIleArgTyrglyLeuValThrAsnGluIleAlaMetValGlnAlaArgGly
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      63  AATATTGTTATTCGTTATGCGCTTGTCACGAAAGAGATAGCATGTCACAGCCGGGT
      122

QY      822  ArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIle
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      123  CGAGCCAGAGCTGATGAAAGCAGCATGTCCTGTCACCAAGAGTGGCTCAGAGTTACC
      182

QY      842  GluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysValAlaIleIscysVal
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      183  GAACGGAGATTGTTAATGATTTCCGAGAGAGATGATGATTAAGCTATTAAACCGTGT
      242

QY      862  GlnAsnMetLysProGluGluTyrAlaHisIleLysIleLeuGluLeuGlnMetGlnSerIle
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      243  CAAACATGAAACCAAGAGATGACACATAGATTGGAATTGACAGGTGCAAGTATC
      302

QY      882  MetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSer
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      303  CTGGAAAAGAAAATGAAGTCGAAAGAAAGCATTTGCAAGCAATATACCAATCCATCG
      362

QY      902  LeuIleThrPheLeuCybLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHis
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      363  TTAATAACACTTCTCTGCAAAATTTGTAGCATGCTGCTGCTCGGAGAAAACATCCAT
      422

QY      922  ValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleVal
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      423  GTCAATTGAGAGATGATCATGTCAATATGACACCAAGATTCAAGGACTCTACATTGTA
      482

QY      942  ArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIle
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      483  AGAGAAAACMAAGCACTGCAAAAAGAAATTTGCTGATTATCAGACCAATGAGAGATTATC
      542

QY      962  CysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuProCys
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      543  TGCAAGTGTGGCAGGCTTGGGGAACAATGATGTCACAAAGGTTAGATTGCTTGT
      602

QY      982  LeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLys
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      603  CTTAAATAAGGAATTTTGTAGTCAATTTCAAAAATAACTCACCGAAGAAACAGTCAAG
      662

QY      1002  LysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGlyCysCysLeuPhe
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      663  AAGTGGGTGAATTCCTATCAGATTTCTCTGATCTTGACTACTCAGATACTGCTGTAT
      722

```

Qy	1022	SerAspGluAsp	1025
Db	723	AGTGATGAAGAT	734

Search completed: March 13, 2006, 21:29:16  
Job time : 9431 secs

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 12, 2006, 04:58:35 ; Search time 688 Seconds  
(without alignments)  
2648.257 Million cell updates/sec

Title: US-09-515-363C-2  
Perfect score: 5311  
Sequence: 1 MSGNGYSTDENFRYLISCPRFRA.....LPITFPNLDYSECCLFSDDED 1025

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 segs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/absas/ABSSWEB.spool/US09515363/runat\_10032006\_181133\_8032/app\_query.fasta\_1  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blotsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs802p -USER=US09515363 @CGN\_1\_1\_512 @runat\_10032006\_181133\_8032  
-NCPU=6 -ICPU=3 -NO\_MMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq:\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq:\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq:\*  
9: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2072	39.0	1392	3	US-09-023-655-48 Sequence 48, Appli
2	1352	25.5	2590	3	US-09-799-451-771 Sequence 771, App
3	498	9.4	458	3	US-09-907-907A-32 Sequence 32, Appli
4	472	8.9	301	2	US-08-143-576-7 Sequence 7, Appli
5	472	8.9	301	3	US-09-221-268D-12 Sequence 12, Appli
6	406	7.6	6184	3	US-09-590-968B-1 Sequence 1, Appli
7	392.5	7.4	1664976	3	US-08-916-421B-1 Sequence 1, Appli
8	392.5	7.4	1664976	3	US-09-692-570-1 Sequence 1, Appli
9	335.5	6.3	7037	3	US-09-853-768-3 Sequence 3, Appli

10	289.5	5.5	1512	3	US-09-408-020-65	Sequence 65, Appli
11	289.5	5.5	32998	3	US-09-408-020-1	Sequence 1, Appli
12	284	5.3	1509	3	US-09-408-020-33	Sequence 33, Appli
13	284	5.3	42432	3	US-09-408-020-2	Sequence 2, Appli
14	283.5	5.3	2354	3	US-10-104-047-1641	Sequence 1641, Ap
15	227.5	4.3	193	2	US-08-143-576-6	Sequence 6, Appli
16	227.5	4.3	193	3	US-09-221-268D-11	Sequence 11, Appli
17	217.5	4.1	8257	3	US-09-595-684B-30	Sequence 30, Appli
18	212	4.0	5852	3	US-09-853-768-10	Sequence 10, Appli
19	211.5	4.0	8503	3	US-09-620-312D-130	Sequence 130, App
20	207.5	3.9	6773	3	US-09-166-350-27	Sequence 27, Appli
21	200	3.8	12214	3	US-09-949-016-17284	Sequence 17284, A
22	199.5	3.8	4868	2	US-08-139-937-12	Sequence 12, Appli
23	199.5	3.8	4868	6	PCT-US93-11310-12	Sequence 12, Appli
24	195.5	3.7	4407	3	US-09-949-016-1690	Sequence 1690, Ap
25	195	3.7	10300	3	US-09-949-016-636	Sequence 636, App
26	194.5	3.7	8789	2	US-08-328-254-5	Sequence 5, Appli
27	193.5	3.6	1986	3	US-09-107-532A-23	Sequence 23, Appli
28	193.5	3.6	8590	3	US-09-949-016-5562	Sequence 5562, Ap
29	191.5	3.6	10136	2	US-08-353-700-2	Sequence 2, Appli
30	191.5	3.6	10136	6	PCT-US95-16216-2	Sequence 2, Appli
31	190.5	3.6	38575	3	US-09-949-016-17304	Sequence 17304, A
32	190.5	3.6	119153	3	US-09-949-016-12378	Sequence 12378, A
33	188.5	3.5	4363	2	US-08-685-576-5	Sequence 5, Appli
34	188.5	3.5	4848	3	US-09-976-594-295	Sequence 295, App
35	187.5	3.5	6921	3	US-09-643-597-117	Sequence 117, App
36	187.5	3.5	6921	3	US-09-480-884A-117	Sequence 117, App
37	187.5	3.5	6921	3	US-09-542-615A-117	Sequence 117, App
38	187.5	3.5	6921	3	US-09-606-421B-117	Sequence 117, App
39	187.5	3.5	6921	3	US-09-221-107-117	Sequence 117, App
40	187.5	3.5	6921	3	US-09-466-396A-117	Sequence 117, App
41	187.5	3.5	6921	3	US-09-476-496A-117	Sequence 117, App
42	187.5	3.5	6921	3	US-09-630-940B-117	Sequence 117, App
43	187.5	3.5	6921	3	US-09-285-479-117	Sequence 117, App
44	187.5	3.5	6921	3	US-10-007-700-117	Sequence 117, App
45	187.5	3.5	7045	3	US-09-919-172-28	Sequence 28, Appli

ALIGNMENTS

RESULT 1  
US-09-023-655-48  
Sequence 48, Application US/09023655  
Patent No. 6607879  
GENERAL INFORMATION:  
APPLICANT: Cocks, Benjamin G.  
APPLICANT: Susan G. Stuart  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023, 655  
FILING DATE: HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1392 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: HMC1NOT01  
CLONE: 004700  
US-09-023-655-48

Alignment Scores:  
Pred. No.: 2.35e-219 Length: 1392  
Score: 2072.00 Matches: 400  
Percent Similarity: 94.8% Conservative: 0  
Best Local Similarity: 94.8% Mismatches: 22  
Query Match: 39.0% Indels: 0  
DB: 3 Gaps: 0

US-09-515-363C-2 (1-1025) x US-09-023-655-48 (1-1392)

QY 604 LeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyr 623  
DB 3 TTGAGGAAGTACATGAGGCCCTACAATTAATGACACATTCGAATGATAGTGCCTAT 62  
QY 624 ThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAlaValIleGluAsp 643  
DB 63 ACTCATCTTGAAACTTCTATAATGAAGAGAAAGATAGAAAGTTGCAGTCATANNNNNN 122  
QY 644 AspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeu 663  
DB 123 NNTTA 182  
QY 664 LysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPheGluAsn 683  
DB 183 AAGAAACCTTGAAACTGAGTGAACAGATAGATTTCATGACTTATTTTGAAGAAC 242  
QY 684 AsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLys 703  
DB 243 AATAAATGTTGAAGAGCTGGCTGAAGAACCCAGAAATGAAATGAAGAGCTGACCAA 302  
QY 704 LeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIle 723  
DB 303 TTAAAGAAATACCATAATGAGCAATATACTAGACTGAGGAATCAGACGAGCAATAATC 362  
QY 724 PheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGluAsnGluLys 743  
DB 363 TTACAAAAACACGACAGAGTCATATGCGCTTCCAGTGATTTACTGAAAAATGAAAA 422  
QY 744 PheAlaGluValGlyValLysAlaHisIleLeuIleGlyAlaGlyHisSerSerGluPhe 763  
DB 423 TTTGCTGAAGTAGAGTCAAGGCCCAACCATCTGATTGAGCTGGACACAGCAGTAGATT 482  
QY 764 LysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIle 783  
DB 483 AAACCCATGACACAGATGAACAAAAAGAGTCATTAGTAATTTGCACTGGAATAATA 542  
QY 784 AsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIle 803  
DB 543 AATCTGCTTATTCGCTACACACAGTGCAGAGAGCTGTGATATTAAAGATGTAACATT 602  
QY 804 ValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAla 823  
DB 603 GTTATCCGTTATGCTCTGCTACCAATGAATAAGCCATGCTCCAGGCCCGTGTGAGCC 662  
QY 824 ArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHis 843  
DB 663 AGAGCTGATGAGAGCACCTACGCTCTGCTGCTCACAGTGTTCAGAGATTATGAAACGT 722

QY 844 GluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsn 863  
DB 723 GAGACAGTTAATGATTTCCAGAGAGAGATGATGATAAAGCTATATGTTGTTCAAAAT 782  
QY 864 MetLysProGluGluTyrAlaHisIleLysIleLeuGluLeuGlnMetGlnSerIleMetGlu 883  
DB 783 ATGAACACGAGAGAGTATGCTCATTAAGATTTTGAATTACAGATGCAAGTATAATGAA 842  
QY 884 LysLysMetLysThrLysArgAsnIleAlaIleHisIleTyrLysAsnAsnProSerLeuIle 903  
DB 843 AAGAAATGAAAAACAGAGAATATTTGCCAAGCATTTACAAGATAACCATCACTAATA 902  
QY 904 ThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIle 923  
DB 903 ACTTTCCTTGCAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATT 962  
QY 924 GluLysMetHisIleValAsnMetThrProGluPheLysGluLeuTyrIleValArgGlu 943  
DB 963 GAGAAATGCATCAGTCATATATGACCCAGAAATCAAGAACTTTCATTGTAAGAGAA 1022  
QY 944 AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLys 963  
DB 1023 AACAAAGCACTGCAAAAGAGTGTGCCGACTATCAAAATAAATGATGTAATCATCTGCAA 1082  
QY 964 CysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLys 983  
DB 1083 TGTCGCCAGGCTTGGGGAACAATGATGTGTGCACAAAGCTTAGATTGCTGTCTCAA 1142  
QY 984 IleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTyr 1003  
DB 1143 ATAAGAAATTTGTAGTGTGTTTCAAAATATATTCAACAAAGAAACATACAAAAAAGTGG 1202  
QY 1004 ValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPheSerAsp 1023  
DB 1203 GTAGAAATTAACCTATCACATTTCCTCAATCTTGACTATTCAGAATGCTGTTATTAGTGAT 1262  
QY 1024 GluAsp 1025  
DB 1263 GAGGAT 1268  
RESULT 2  
US-09-799-451-771  
Sequence 771, Application US/09799451  
Patent No. 6783969  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Zhou, Ping  
APPLICANT: Goodrich, Ryle  
APPLICANT: Asundi, Vinod  
APPLICANT: Ren, Feiyan  
APPLICANT: Zhang, Jie  
APPLICANT: Xue, Aidong J.  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Ma, Yungqing  
APPLICANT: Yamazaki, Victoria  
APPLICANT: Chen, Rui-hong  
APPLICANT: Wang, Zhiwei  
APPLICANT: Wang, Dunrui  
APPLICANT: Yang, Yonghong  
APPLICANT: Wehrman, Tom  
APPLICANT: Ghosh, Reena  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6783969el Nucleic Acids and  
FILE REFERENCE: Polypeptides  
CURRENT APPLICATION NUMBER: US/09/799,451  
CURRENT FILING DATE: 2001-03-05  
NUMBER OF SEQ ID NOS: 948  
SOFTWARE: pt\_fl\_genes Version 2.0  
SEQ ID NO 771  
LENGTH: 2590

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (190)..(2223)
US-09-799-451-771

Alignment Scores:
Pred. No.: 5.17e-139 Length: 2590
Score: 1352.00 Matches: 292
Percent Similarity: 59.7% Conservative: 136
Best Local Similarity: 40.7% Mismatches: 237
Query Match: 25.5% Indels: 52
DB: 3 Gaps: 12

US-09-515-363C-2 (1-1025) x US-09-799-451-771 (1-2590)

Qy 306 LeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAsnIle 325
Db 190 ATGAGCTTCGGTCTACCAATGGAGGTGATCATGCTGCCCTGGAGGGCAAGATATC 249
Qy 326 IleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIleAlaLysAsp 345
Db 250 ATCATCTGGCTGCCACCGGGTGGCGGAAGACCCGGGGCGCTGCTTATGTGGCCAAAGCGG 309
Qy 346 HisLeuAspLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLys 365
Db 310 CACCTAGAG-----ACTGTGATGAGCCAAAGGTGTGTATGTGTCAACAGG 357
Qy 366 ValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeuLysTrpTyr 385
Db 358 GTGCACCTGTGTACCAG---CATGTGAAGAGTTCAAGCGCATGCTGATGGACGCTGG 414
Qy 386 ArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPheProGluValLys 405
Db 415 ACCGTGACAACCTGAGTGGGACATGGACCACGTGCTGGCTTGGCCACTGGCCCGG 474
Qy 406 SerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGlu 425
Db 475 TGCATGACCTGTCTCATCTGCACAGCAGAGCTTCTGCAGATGGCACTGACCAACCCCGAG 534
Qy 426 AsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIleIleAspGluCys 445
Db 535 -----GAGGAGGAGCACGTGAGCTCACTGTTCTTCCTGATCGTGTGATGAGTGC 588
Qy 446 HisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGlnLys 465
Db 589 CACCACACGACACAAGACACCGTCTTACAACGTCATCATGAGCCAGTACCTAGAACTTAA 648
Qy 466 IleuLysAsnAsnArgLeuLysLysGluAsnLysProValIleProLeuProGlnIleLeu 485
Db 649 CTCACAGAGGGCACAG-----CCGCTACCCACAGGTGCTG 681
Qy 486 GlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAlaLysAlaGluHis 505
Db 682 GGTCTCACAGCCTCCCAAGCACTGGCGGGGCTCCAAACTCGATGGGGCCATCAACCAAC 741
Qy 506 IleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLeuThrValLysGluAsnLeu 525
Db 742 GTCTTGACGCTGTGTCCAACTTGACACGTTGTCATCATGTCAACCCCAAGACTGCTGC 801
Qy 526 AspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThr 545
Db 802 CCCACGCTGCAGAGACACAGCCAAGCCTTGCAACAGTCAACCTGTGCCACAGCGC 861
Qy 546 ArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyrCys 565
Db 862 AGCCAGGATCCGTTTGGGACTTGTGTAAGAAGCTCATGGAACCAATCCATGACACACTG 921
Qy 566 GlnMetSerPrometSer---AspPheGlyThrGlnProTyrGluGlnTrpAlaIleGln 584
Db 922 GAGATGCCCTGAGTTGAGCCGGAATTTGGGACCGCAATGTATGAGCAGCAGGTGTGAG 981

Qy 585 MetGluLysLysAlaAlaLysGlyAsnArgLysGluArgValCysAlaGluHisLeu 604
Db 982 CTGAGTGAAGGCTGCGGCTTTGGCTGGCTTCAAGAGCAACGGGTGTATGCGCTTCACTG 1041
Qy 605 ArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyrThr 624
Db 1042 AGCGCTACAATGACGCGCTCATCATCATGACACCGTCCGCGCGGTGATGCTTGGCT 1101
Qy 625 HisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspAsp 644
Db 1102 GCGGTGACAGATTCTTATCACAGGAGCAGCAGTCACTAAACCAAGATCCTG----- 1152
Qy 645 SerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspAspLeuLys 664
Db 1153 -----TGT----- 1155
Qy 665 LysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsn 684
Db 1156 -----GCCGAGCGCGGCTGTGGCCCTGTGTGATGACCGCAAG 1194
Qy 685 LysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeu 704
Db 1195 AATGACTGGCCCACTTGGCACTCATGAGCCCA---GAGAATCCAAACTGAGATGCTG 1251
Qy 705 ArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePhe 724
Db 1252 GAAGAATCCTGCAAAGGAGATTCAGT--AGCTTAACAGCCCTCGGGGTATCATCTTC 1308
Qy 725 ThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPhe 744
Db 1309 ACCCGCAACCCGCCAAGCGCACACTCCCTCCTGCTGTGGCTCCAGCAGCAGCAGGCGCTG 1368
Qy 745 AlaGluValGlyValLysAlaHisIleLeuIleGlyAlaGlyHisSerSerGluPheLys 764
Db 1369 CAGACTGTGACATCCGGGCCAGCTACTGATGGGGCTGGAAACAGCAGCAGACAGCACC 1428
Qy 765 ProMetThrGlnAsnGlnLysGluValIleSerLysPheArgThrGlyLysIleAsn 784
Db 1429 CACATGACCCAGAGGAGCAGCAAGAAGTGATCCAGAAGTTCCAAAGATGGAACCTGAAC 1488
Qy 785 LeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleVal 804
Db 1489 CTTGTGTTGCCACGAGTGTGGCGGAGAGGCGGTGGACATCCACATGTCAATGTGCTG 1548
Qy 805 IleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArg 824
Db 1549 GTGCGTTATGGGCTCTTGACCAATGAATCTCCATGCTCCAGGCCAGGGGCGTCCCGG 1608
Qy 825 AlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHisGlu 844
Db 1609 GCCATCAGAGTGTATACCGCTTTGTAGCAACTGAAGGTAGCCGGGAGCTGAAGCGGAG 1668
Qy 845 ThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsnMet 864
Db 1669 CTGATCAACGAGCGCTGAGACGCTGATGAGCAGCAGTGGCTGTGCAAGAAATG 1728
Qy 865 LysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGluLys 884
Db 1729 GACCAGCGCGAGTACCAAGGCCAAGATCCGGGATCTGCAGCAGCAGCCTTGACCAAGCGG 1788
Qy 885 LysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIleThr 904
Db 1789 GCGGCCCCAGGACGCCAGCGGGAAGAACCAAGCGGACAGTTCACAGTGAACAGTGCAG 1848
Qy 905 PheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIleGlu 924
Db 1849 CTACTCTGCATCAACTGCATGTGTGCTGTGGCCATGGCAGCGACCTGGCGAAGGTGAG 1908
Qy 925 LysMetHisHisValAsnMetThrProGluPheLysGluLeuTyr---IleValArgGlu 943
Db 1909 GGCACCCACCATGTCAATGTGAACCCCAACTTCTGAACTACTATATGTCTCCAGGAGT 1968
Qy 944 AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyLysIleLeuLys 963

Db 1969 CCTGTGTCATCAACAAGCTTCAAGAGCTGGAAGCCCTGGGGTGTCTACAGCTGCAGG 2028  
QY 964 ---CysGlyGlnAlaTrpGlyThrMetMetValHisGlyGlyLeuAspLeuProCysLeu 982  
Db 2029 AACTGTGGGAGGCTCTGGGGCTCTGCAGATGATCTACAAGTCAAGTGAAGCTGCCAGTGCTC 2088  
QY 983 LysIleArgAsnPhenValValValPheLysAsnAsnSerThrLysLysGlnTyrLysLys 1002  
Db 2089 AAAGTCCGCAGC-----ATGCTGCTGGAGAGACCCCTCAGGGGGCGGATCCAGCCAAAAAG 2142  
QY 1003 TrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019  
Db 2143 TGGTCCCGCGCTGCTTCTCCGTGCTGACTTGTGACTTCTCTGCAGCATTTGT 2193

RESULT 3

US-09-907-907A-32  
; Sequence 32, Application US/09907907A  
; Patent No. 6951923  
; GENERAL INFORMATION:  
; APPLICANT: Fisher, Paul B.  
; APPLICANT: Leszczyniecka, Magdalena  
; TITLE OF INVENTION: GENES DISPLAYING ENHANCED EXPRESSION DURING CELLULAR SENESENCE A  
; TITLE OF INVENTION: TERMINAL CELL DIFFERENTIATION AND USES THEREOF  
; FILE REFERENCE: A34584-A-PCT-USA (070050.1664)  
; CURRENT APPLICATION NUMBER: US/09/907,907A  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: US 09/243,277  
; PRIOR FILING DATE: 1999-02-02  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 32  
; LENGTH: 458  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 342, 355, 365, 368, 375, 381, 385, 414, 445  
; OTHER INFORMATION: a or c or g or t  
US-09-907-907A-32

Alignment Scores:  
Pred. No.: 1.43e-45 Length: 458  
Score: 498.00 Matches: 90  
Percent Similarity: 98.9% Conservative: 0  
Best Local Similarity: 98.9% Mismatches: 1  
Query Match: 9.4% Indels: 0  
DB: 3 Gaps: 0

US-09-515-363C-2 (1-1025) x US-09-907-907A-32 (1-458)

QY 935 PheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyr 954  
Db 3 TTCAAGAACTTTACATTGTAAGAGAAAAACAACACTGCAAAAGAAGTGCCTAT 62  
QY 955 GlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHis 974  
Db 63 CAATTAATGCTGAATCATCTGCAAAATGTGGCCAGGCTTGGGAACAATGATGTCAC 122  
QY 975 LysGlyLeuAspLeuProCysLeuLysIleArgAsnPhenValValPheLysAsnAsn 994  
Db 123 AAAGCTTAGATTGCTTGTCTCAAAATTAAGGAATTTGTAGTGTCTTCAAAAATAAT 182  
QY 995 SerThrLysLysGlnTyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAsp 1014  
Db 183 TCACAAAGAAACAATACAAAGAGTGGGTAGAAATTACCTATCACATTTCCCAATCTTGAC 242  
QY 1015 TyrSerGluCysCysLeuPheSerAspGluAsp 1025  
Db 243 TATTCAGAATGCTGTTATTATTAGTATGATGAGAT 275

RESULT 4  
US-08-143-576-7

; Sequence 7, Application US/08143576  
; Patent No. 5643761  
; GENERAL INFORMATION:  
; APPLICANT: Fisher, Paul B.  
; APPLICANT: Jiang, Hongping  
; TITLE OF INVENTION: METHOD FOR GENERATING A SUBTRACTED  
; TITLE OF INVENTION: CDNA LIBRARY AND USES OF THE GENERATED LIBRARY  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White, c/o Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/143,576  
; FILING DATE: 25-OCT-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 43563/JPW/AKC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 977-9550  
; TELEFAX: (212) 664-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 301 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-143-576-7

Alignment Scores:  
Pred. No.: 5.3e-43 Length: 301  
Score: 472.00 Matches: 97  
Percent Similarity: 98.0% Conservative: 0  
Best Local Similarity: 98.0% Mismatches: 2  
Query Match: 8.9% Indels: 1  
DB: 2 Gaps: 0

US-09-515-363C-2 (1-1025) x US-08-143-576-7 (1-301)

QY 476 LysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGly 495  
Db 3 AAACCACTGATTCCTCTCAGATACTGGGACTTAACAGCTTCACTGCTGTTGAGGG 62  
QY 496 AlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAla 515  
Db 63 GCCACGAAGCAAGCCAAAGCTGAAGAACAACATTTAAACTATGTGCTATCTTGATGCA 122  
QY 516 PheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluPro 535  
Db 123 TTTACTATTAAACTGTTAAAGAAAACTTGATCACTGAAAAACCAAAATACAGAGCA- 181  
QY 536 CysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGlyLysLeuLeu 555  
Db 182 TGCAAGAAAGTTGCAATGAGATGCAACACAGAGATCCATTAAAGAGAACTTCTA 241  
QY 556 GluIleMetThrArgIleGlnThrTyrCysGlnMetSerPrometSerAspPheGly 574  
Db 242 GAAATATGACAAAGATTAACCTTATTGTCAATGAGTCCAATGTCAGATTTTGA 298

```
RESULT 5
US-09-221-268D-12
; Sequence 12, Application US/09221268D
; Patent No. 6720408
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH SPECIFICALLY HYBRIDIZE WITH
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE ENCODING MDA-7 AND PHARMACEUTICAL COMPOSIT
; FILE REFERENCE: A34534-A-A (070050.1637)
; CURRENT APPLICATION NUMBER: US/09/221,268D
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 08/316,537
; PRIOR FILING DATE: 1994-09-30
; PRIOR APPLICATION NUMBER: 08/143,576
; PRIOR FILING DATE: 1993-10-27
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 301
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-221-268D-12

Alignment Scores:
Pred. No.: 5.3e-43 Length: 301
Score: 472.00 Matches: 97
Percent Similarity: 98.0% Conservative: 0
Best Local Similarity: 98.0% Mismatches: 2
Query Match: 8.9% Indels: 1
DB: 3 Gaps: 0

US-09-515-363C-2 (1-1025) x US-09-221-268D-12 (1-301)
QY 476 LysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGly 495
Db 3 AAACCAAGTGAATCCCTTCCTCAGATACTGGACTTAACAGCTTCACCTGGTGGAGGG 62
QY 496 AlaThrLysGlnAlaLysAlaGluHisIleLeuLysLeuCysAlaAsnLeuAspAla 515
Db 63 GCCACGAAGCAAGCCAAAGCTGAAGACACATTTTAAACTATGTGCTTATCTTGATGCA 122
QY 516 PheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnIlePro 535
Db 123 TTTACTATTAAACTGTCTTAAAGAAACCTTGATCAACTGAAAAACCAATACAGAGCA- 181
QY 536 CysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeu 555
Db 182 TGCAGAGAGTTTGCATTGCAGATGCAACAGAGAGATTCATTAAAGAGAACTTCTA 241
QY 556 GluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGly 574
Db 242 GAAATATGACAAAGATTCAAACTTATTGTCAATGAGTCCAATGTCAGATTTTGA 298

RESULT 6
US-09-590-968B-1
; Sequence 1, Application US/09590968B
; Patent No. 6737561
; GENERAL INFORMATION:
; APPLICANT: Ray, Animesh
; APPLICANT: Golden, Teresa Ann
; TITLE OF INVENTION: GENE ENCODING SHORT INTEGUMENTS AND USES THEREOF
; FILE REFERENCE: 176/60581
; CURRENT APPLICATION NUMBER: US/09/590,968B
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/138,316
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6184
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
```

```
US-09-590-968B-1

Alignment Scores:
Pred. No.: 2.06e-33 Length: 6184
Score: 406.00 Matches: 231
Percent Similarity: 39.8% Conservative: 158
Best Local Similarity: 23.6% Mismatches: 316
Query Match: 7.6% Indels: 275
DB: 3 Gaps: 47

US-09-515-363C-2 (1-1025) x US-09-590-968B-1 (1-6184)
QY 20 AlaArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuPro 39
Db 127 TCTAGGGTTTGTCTGCTCTCTCTCGTACCC- - - - -TTTTTACCT 168
QY 40 Ala-----GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGln 57
Db 169 GCAACAACAACACTTCAAAATTGGCGGTGTTGCTACGCTCTA-TCTAACCTAATCTGTCA 227
QY 58 AlaValGluLeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArg 77
Db 228 CAAACAACCTCTCTCTCTCAACCCCTTTTCTGGGTTATTCAAT- - - - -TCTCGT 278
QY 78 GluPheValGluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnPro 97
Db 279 GCTTTT- - - - -GTTCTGTTTCTTCTCTGGG 305
QY 98 GluLeuThrAspLeuProSerProSerPheGluAsnAla- - - - -HisAspGlu 113
Db 306 GATTGGTTTCTTG- - -AGTGAGTTTTCCTCTTCTTCTTATGTTCTGATTGATTAT 362
QY 114 TyrLeuGlnLeu- - - - -LeuAsnLeuGlnPro- - - - -ThrLeuVal 126
Db 363 TATATAGAATTATGTATGAGAGATGAGCCTAGAGAAACCAATTAAGCCTTCTTATT 422
QY 127 AspLysLeuLeuValArgAspValLeuAspLysCys--MetGluGluLeuLeuThr 145
Db 423 GGCTAGATGCTTGC-GAGGACATC- - - - -TCTTGATCTTATCGATGATCTCGTCT 475
QY 146 IleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsnGluSerGlyValArg 165
Db 476 GAATTTGATCCTCTCTGTTGCTGTCATGAATCCACT- - -GATGAAAACGGCGTC- - 529
QY 166 GluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAlaPheLeuAsnValLeu 185
Db 530 - - - - -ATCAATGATTTTTCGGTGGATGATGATCATTTTA 565
QY 186 ArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySerAspCysSerGluSer 205
Db 566 GATAGTATCAAGAACGGT- - - - -GGAGCCTTACCACAAACAATGGCGTTCTGATACC 616
QY 206 AsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnValGluGluGlnLeu 225
Db 617 AATTCTCAAAATCAACGAGGTTACT- - - - -CCTCAGGTT- - - - -ATTGCT 661
QY 226 SerThrThrValGlnProAsn--LeuGluLysGluValTrpGlyMetGluAsnSer 244
Db 662 AAGGACAGTGAAGAGAAATGGGTTGCAAAAAGAAATGGCGGTAAGAGAGAGCAATTCTCG 721
QY 245 SerGluSerSerPheAlaAsp- - - - -SerSerValValSerGluSer 258
Db 722 AAAGAGAGAGAGACAGATAGAGAGAGAGCTAGGGTTTGTAGTTATCAGAGTGAAGG 781
QY 259 AspThrSerLeuAlaGluGlySerValSerCysLeuAspGlu- - - - - 272
Db 782 AGTAACCTTTCAGGTAGAGGGCATGTTAATAATTCTACGAGGAGAGATAGTTTATGAAT 841
QY 273 - - - - -SerLeuGlyHisAsn- - - - - 277
Db 842 AGGAACGTACTCTGTAATTGGGACGAGGGGGTAACAATTAAGAGAAAGGCAATGTAAAC 901
QY 278 - - - - -SerAsnMet 280
```

```
Db      902 AATTACAGAGAGATGGTAGAGATAGAGAAGTTAGGGCTTATTGGAGAGCGATAAAGTT   ::
Qy      281 GlySerAsp-----SerglyThrMetGlySerAspSerAsp----- 292
Db      962 GGTTCCAATGAGTTGGTTATAGGTCAGGAGCTTGGGAAGCTGATCATGAAGAAGATGTT 1021
Qy      293 -----GluGluAsnValAlaAla 298
Db      1022 AAGAAAGTAGTGGTGAAACCCGGAATGCGATGTCAAGCGACAGAGAGAAC----- 1072
Qy      299 ArgAlaSerProGluPro-----GluLeuGlnLeuArgProTyrGln 312
Db      1073 AAGAGTAAGCCTGAAGAACGTTAAAGAGAGAGAGGTTGTGAAGACCAAGCAGCGATACCA 1132
Qy      313 MetGluValAlaGlnProAlaLeuGluGlyLysAsnIleIleIleCysLeuProThrGly 332
Db      1133 TTGGATGTTCTTGAAACAAAGCT--AAAGCGAAAAACAGATTGCTTCTTGAGACCGGT 1189
Qy      333 SerGlyLysThrArgValAlaValTyr-----IleAlaLysAspHisLeuAsp 348
Db      1190 GCTGGAAGAAGACACTTATCGCATTTCTTATTAAAGTGTTCATAAGGATCTGATGAGC 1249
Qy      349 LysLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLysValLeuLeu 368
Db      1250 CAGAAACAGAAAATGCTCTCG-----GTGTTCTTGGTCCCAAGTGCCTTGTG 1297
Qy      369 ValGluGlnLeuPheArgLysGlu-----PheGlnProPheLeuLysLys 383
Db      1298 GTTTATTCAGCA-AGCAGAAAGTATCCGTAATCAAACTTGTTTCAAGTTGGACATTATTG 1356
Qy      384 TrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-LysIleSerPheProGluVal 403
Db      1357 TGGTGAGATGGGACAGAGACTTTTGGGATTCTCGAAGGTGGCAACGAGAGTTT----- 1408
Qy      403 IValIysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAs 423
Db      1409 ----GAGTCTAAGCAGGTTCTAGTTATGACAGACACAATTTCTGTGAATATACTG----- 1459
Qy      423 nLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIleIleLeAs 443
Db      1460 -----AGACACAGATCATTTAGATGGAACAATTGATCTTCTTATCTCGA 1506
Qy      443 pGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMe 463
Db      1507 CGAGTGTCAACACGCTGTCAAGAAACATCCATACTCTTTAGTGATGTCAAGAGTTTACCA 1566
Qy      463 tGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIleProLeuProGl 483
Db      1567 TACAACCTCTAAAGATAAAGA-----CCTGC 1593
Qy      483 nIleLeuGlyLeuThrAlaSerPro--GlyValGlyGlyAlaThrLysGlnAlaLysAl 502
Db      1594 CATCTTTGGAATGACTGCTCGCTGTTAATTAAAGGTGTTCAGGCCAAGTAGATTG 1653
Qy      502 agLnuGlnHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThrValLy 522
Db      1654 TGGCATTAAGATACGTAACCTCGAGACCAAGTTGGATCT--ACGGTTGTACTATAAA 1710
Qy      522 sGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAlaIleAl 542
Db      1711 AGATCGA---AAAGATTAGAGAAACATGTGCCTATGCCTTCAGAGATAGTCGTGAGTA 1767
Qy      542 asp---AlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArgIl 561
Db      1768 TGACAAAGCTGCTACTATGTGCTCTTCATGAGACAATAAAGCAATGATTGACGCTGT 1827
Qy      561 eGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGlnTr 581
Db      1828 TGAAGAAGCGGCACAAAGCAAGTTCAAGGAAA-----AGCAAGTG 1866
Qy      581 pAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCysAl 601
```

```
Db      1867 GCAATTTATGGGGCTAGGGATGCTGAGACAAAGATGAATTGAGACAGTTTATGGCGT 1926
Qy      601 agLnuHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAs 621
Db      1927 CTCTGAAAAGAACCGAGAGCGGATGCTGCTGCCAATTGATTGCATTAACCTAGAGCTATCAA 1986
Qy      621 p-----AlaTyrThrHisLeuGlnThrPhe----- 629
Db      1987 TTATACTCTTGCTGAATTGGGTCAATGCTGTGTCTTACAAAGGTGGACAAATCATCTTGTG 2046
Qy      630 -----TyrAsnGluGluLysAspLysPhe----- 638
Db      2047 TGTCTTGCAAAAGTAGAGAGAGGGTGAAATTTCCAAGTCGACGTGAAGTTTCAAGAAATCATA 2106
Qy      639 -----AlaValIleGlu-- 642
Db      2107 CCTCAGTGAGGTGGTGTCATCTTGCAATGTGAGCTTCTGGAAGCGCTGCTGTAATA 2166
Qy      643 -----AspAspSerAspGluGl 648
Db      2167 AGTCGGCGCGGAGAGTTGGCAAAACGAAATGTGAATGCACATGACGAGATGAGAGAGG 2226
Qy      648 Y-----GlyAspAspGluTyrCysAspGlyAspGluAspGluAspAspLeuLysLysPr 666
Db      2227 AGAGCTCCCTGATGATCTCTGTGCTCGGAGAGGAGACACGTTGATGAATAAGCGGC 2286
Qy      666 oleuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysBme 686
Db      2287 CGCAGTGGCTGAT-----GGGAAAGT 2307
Qy      686 tLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAs 706
Db      2308 TACTCCAAAAGTA-----CAATCATGTATCAAACTA----- 2338
Qy      706 nThrIleMetGluGlnTyrThrArgThrGluGlnSerAlaArgGlyIleIlePheThrLy 726
Db      2339 -----CTCCTCAAAATATCAGCACACAGCTGAT--TTTCGAGCTATGTTTTCGTTGA 2388
Qy      726 sThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGl 746
Db      2389 GAGGGTGGTTGCTGCTTGGTCTTCT--AGGTTTTCGCGA 2430
Qy      746 uValGly-----ValLysAlaHisLeuIleGlyAlaGlyHisSerSergl 762
Db      2431 GCTGCTTCGCTTAGTTTATACGGTGTGCCAGCATGATTGGACACAAATTAACGCCAGGA 2490
Qy      762 uPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLy 782
Db      2491 GATGAAA-----TCAATCTCAAAATGCAGAGATCAAAATTTCCAAATTCGAGATGGGCA 2541
Qy      782 sIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAs 802
Db      2542 TGTGACACTGTAGTGGCCACAAGCGTTGCTGAGGAAGGACTTGATATTAGGCAATGTAA 2601
Qy      802 nIleValIleArgTyrGlyLeuValThrAsnGlnIleAlaMetValGlnAlaArgGlyAr 822
Db      2602 CGTTGTTATGCGTTTGCACCTTGCAAGAAGACGGTGTGGCATATACATTAGTCTCGTGGCG 2661
Qy      822 gAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySer 838
Db      2662 GGCAGAAAGCCTGGATCAGACTACATACTCATGTGTTGAGAAGAGAAAT 2710

RESULT 7
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
```

;; CURRENT FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: US 60/024,428  
;; PRIOR FILING DATE: 1996-08-22  
;; NUMBER OF SEQ ID NOS: 3  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 1  
;; LENGTH: 1664976  
;; TYPE: DNA  
;; ORGANISM: Methanococcus jannaschii  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (28222)..(28222)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (28257)..(28258)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (84773)..(84773)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (84808)..(84808)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (84812)..(84812)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (98120)..(98120)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (98159)..(98159)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (98239)..(98239)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (98266)..(98266)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (98343)..(98343)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (103998)..(103998)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (148948)..(148948)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (163385)..(163385)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (191989)..(191989)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (191995)..(191995)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (231980)..(231980)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (234187)..(234187)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (234220)..(234220)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (234814)..(234814)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (309398)..(309398)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (309418)..(309418)  
;; OTHER INFORMATION: n equals a, t, c, or g

;; NAME/KEY: misc feature  
;; LOCATION: (312837)..(312837)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (312993)..(312993)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (319226)..(319226)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (559167)..(559167)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (559241)..(559241)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (600992)..(600992)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (622708)..(622708)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (657081)..(657081)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (657203)..(657203)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (674435)..(674435)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (682442)..(682442)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (713652)..(713652)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (741684)..(741684)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (779455)..(779455)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (779676)..(779676)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (855539)..(855539)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (871619)..(871619)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (1084830)..(1084830)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (1096846)..(1096846)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (1119881)..(1119881)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (1130881)..(1130881)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (1310988)..(1310988)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (1313224)..(1313224)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature  
;; LOCATION: (1349473)..(1349473)  
;; OTHER INFORMATION: n equals a, t, c, or g  
;; NAME/KEY: misc feature

```

; LOCATION: (1349491) .. (1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091) .. (1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020) .. (1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912) .. (1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734) .. (1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998) .. (1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854) .. (1664855)
; OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

Alignment Scores:

Pred. No.:	1.03e-27	Length:	1664976
Score:	392.50	Matches:	169
Percent Similarity:	40.1%	Conservative:	125
Best Local Similarity:	23.0%	Mismatches:	241
Query Match:	7.4%	Indels:	199
DB:	3	Gaps:	27

US-09-515-363C-2 (1-1025) x US-08-916-421B-1 (1-1664976)

```
QY      306 LeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyAsnIle 325
      |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1658788 TTGGAGCGCGAGGTGTATCAGCAGATTATTGCGACGAATGCTTTAAGAAAAAGACATTA 1658847

QY      326 IleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIleAlaLysAsp 345
      :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1658848 TGTGTT--TTATCGACAGGTTTAGGTAAACACAGCTATTGCTATTAGTTATAGCAGGT 1658904

QY      346 HisLeuAspLysLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLys 365
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1658905 ATTTTAACAAAAAAG-----GATGGAAGGTTTAACTTGAAGCCCTTCA 1658949

QY      366 ValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeuLysTrpTyr 385
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1658950 AGACCTTTGGTTGAGCAACACTAACACAGATTAAACAGTTGATGAAGAT 1659009

QY      386 ArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPheProGluValValLys 405
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659010 AAAATAATAGCTTTAACTGGAATAATCCAGCCAAAA--AAGAGAGCTGAACCTATATAA 1659066

QY      406 SerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGlu 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659067 AAAGGAAAAATCTTTATAGCTAACACCAACAGTTATAGAAAAACGATATCATATA----- 1659117

QY      426 AsnGlyGluAspAlaGly--ValGlnLeuSerAspPheSerLeuIleIleIleAspGlu 444
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659118 -----GCTGGAAGAATAATGTGATGAATTATTTATTTGATGAGCTGATGAA 1659165

QY      445 CysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGln 464
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659166 GCCCAACACACACAGAGACCATGCTAT-----GCATTTGTAGCAAAA 1659210

QY      465 LysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIleProLeuProGlnIle 484
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659211 AAATTTAAAGATAAA-----TGTCAATATT 1659234

QY      485 LeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAlaLysAlaGluGlu 504
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659235 TTAGGTTTAACGGCATCTCCAGGT-----TCTGATATTGAT 1659270

QY      505 HisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThrValLysGluAsn 524
```

```

Db      1659271 AAAGTCATGGAATTTGTAAAACTTAGCAATTGACACAGCTTGAAGTGAGAACTGAAGAT 1659330

QY      525 LeuAspGlnLeuLys----- 529
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659331 GATGAGGATGTAAACCATACATTGCTTAAAGTAAACTTATCCCAATTGAATTGATTTA 1659390

QY      530 ---AsnGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThrArgGlu--- 547
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659391 CCCAACGAATTTAAAGAGCGTTAAAA-----TTAATAAATGAAGCTTTTAAAGAGAGA 1659444

QY      548 -----AspProPheLysGluLys 553
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659445 TTAATAATATTAAAGATGCTGAGTTATAAATCCATTGCCGATGTAACAAAAACAGAA 1659504

QY      554 LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerPrometSerAspPhe 573
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659505 CTTATTGAGCTAAATATAAGCTATTTCCTAT----- 1659537

QY      574 GlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysGly 593
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659537 ----- 1659537

QY      594 AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle 613
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659538 -----GATGAAGAAGTGAAGTATGAACCTTATAAAAGTTTGTTCAGAGCCTTTAAAACTT 1659591

QY      614 AsnAspThrIleArgMetIleAspAla-----TyrThrHisLeuGlu 627
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659592 ATGCATGCCAAGAAGACTCTTAGAGAGTCAAGGAAGAAGTGTATTTTAAACTATATAAAT 1659651

QY      628 ThrPheTyrAsnGluLysAspLysLysPheAlaValIleGluAspAspSerAspGlu 647
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659652 AAATTTATCCATGCAAAGAACAAATCAGCTAAATCTATTGTTAATGATGAA----- 1659702

QY      648 GlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspAspLeuLysProLeu 667
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659702 ----- 1659702

QY      668 LysLeuAspGluThrAspArgPheLeuMetThrLeuPheGluAsnAsnLysMetLeu 687
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659703 AAAGTTAGAGAGCAGTTAATTATTATGAATAATCAGATGTAGAACAATCCAAAAATTA--- 1659759

QY      688 LysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThr 707
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659760 -----GCTAAAGTTGTGATATGCTTAAAAAT 1659786

QY      708 IleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThr 727
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659787 ATTTTGGAAAA-----AATAAGATGAGAGAAATATATTATCTTGTCTCAATAT 1659834

QY      728 ArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGluAsnGluLysPheAlaGluVal 747
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659835 AGGACACTGTGAGAGAAGATTGTTAATCTTAACTCAAAAT----- 1659876

QY      748 GlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysPrometThr 767
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659877 GGAATTAAGCAATAGAATTATAGACAGGCA--AATAAAGAAAGAAAGGAATGAGT 1659933

QY      768 GlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIle 787
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659934 CAGAAAGACCAATAGAACTATAGAGAGATTTAAAAA--GAGGGAAGTGTTTAGTT 1659990

QY      788 AlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyr 807
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1659991 TCAACAAAGCGTTTCTGAGAGGGAATAGATATTCATCGGTAAATTACATCATATTTTAT 1660050

QY      808 GlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGlu 827
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1660051 GAACCACTGCCATCAGAAATTAGGTTTATTCAGAGAGAGGTAGAGCGATGAGCGGAGAA 1660110

QY      828 -----SerThrTyrValLeuVal-----AlaHisSerGlySer 838
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 1660111 GGAGGGAAGTTTATGTTTAACTAAGGGAACAGCTGATGAAGCTTATTACAGAGT 1660170  
Qy 839 G1yVal1IeGluH1sGluThrValAsnAspPheArgGluLysMetTyrLysAla1le 858  
Db 1660171 GCTTTATACAAAGAAAGGAGATGAGATATTAAATAATATGTTATTGCTAAAT 1660230  
Qy 859 HisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMet 878  
Db 1660231 AAGAGTTACAGAG--AAATTGAAGAAAAATCTAAGAGGAATAAAGAGAGACA 1660287  
Qy 879 GlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLys--- 897  
Db 1660288 GAAGAAATAAAGAAAGAAATGAAATCAAAAACGTGATAAAGAGAACTAAGAGAG 1660347  
Qy 898 -----AsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerVal 912  
Db 1660348 GAAGAGAAAAAACCAAAAAGCCAGTAACGATATTAGATTTCATTAAACAGATTGAAGTT 1660407  
Qy 913 -----LeuAlaCysSerGlyGluAsp----- 919  
Db 1660408 AAGGAAAGGCTTAATCAGAAAGAGATAAAATAAACAGAGATAAAAATTCGAAAAAAG 1660467  
Qy 920 ---IleHisValIleGluLysMetHisHisValAsnMetThrPro----- 933  
Db 1660468 CCAATTAAGATTATTGTAGATGTTAGAGAGAGAAGAAATATGCTTAAGCTTTACATAATTAT 1660527  
Qy 934 -----GluPheLysGluLeu-----TyrIleValArgGluAsnLysAla 946  
Db 1660528 GCAATATATTGAGCTAAATAACATTAGAAAGTGGAGATATTGTTTAAGTGAAGGTAAGTT 1660587  
Qy 947 LeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 1660588 GTTGAGAGAAAGACAGCTGAAGACTTGTAAATTCAATAATT 1660629

## RESULT 8

US-09-692-570-1

: Sequence 1, Application US/09692570

: Patent No. 6797466

: GENERAL INFORMATION:

: APPLICANT: Built et al.

: TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus

: Patent No. 6797466

: TITLE OF INVENTION: jannaschii

: FILE REFERENCE: PB275C1

: CURRENT APPLICATION NUMBER: US/09/692,570

: PRIOR FILING DATE: 2003-01-14

: PRIOR APPLICATION NUMBER: US 60/024,428

: PRIOR FILING DATE: 1996-08-22

: PRIOR APPLICATION NUMBER: US 08/916,421

: NUMBER OF SEQ ID NOS: 20

: SOFTWARE: PatentIn version 3.1

: SEQ ID NO 1

: LENGTH: 1664976

: TYPE: DNA

: ORGANISM: Methanococcus jannaschii

: FEATURE:

: NAME/KEY: misc feature

: LOCATION: (28222)..(28222)

: OTHER INFORMATION: n equals a, t, c, or g

: FEATURE:

: NAME/KEY: misc feature

: LOCATION: (28257)..(28258)

: OTHER INFORMATION: n equals a, t, c, or g

: FEATURE:

: NAME/KEY: misc feature

: LOCATION: (84773)..(84773)

: OTHER INFORMATION: n equals a, t, c, or g

: FEATURE:

: NAME/KEY: misc feature

: LOCATION: (84808)..(84808)

: OTHER INFORMATION: n equals a, t, c, or g

: FEATURE:

: NAME/KEY: misc feature  
: LOCATION: (84812)..(84812)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (98120)..(98120)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (98159)..(98159)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (98239)..(98239)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (98266)..(98266)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (98343)..(98343)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (103998)..(103998)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (148948)..(148948)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (163385)..(163385)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (191989)..(191989)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (191995)..(191995)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (231980)..(231980)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (234187)..(234187)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (234220)..(234220)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (234814)..(234814)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (309398)..(309398)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (309418)..(309418)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature  
: LOCATION: (312837)..(312837)  
: OTHER INFORMATION: n equals a, t, c, or g  
: FEATURE:  
: NAME/KEY: misc feature

LOCATION: (312993) .. (312993)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (319226) .. (319226)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (559167) .. (559167)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (559241) .. (559241)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (600992) .. (600992)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (622708) .. (622708)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (657081) .. (657081)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (657203) .. (657203)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (674435) .. (674435)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (682442) .. (682442)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (713652) .. (713652)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (741684) .. (741684)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (779455) .. (779455)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (779676) .. (779676)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (855539) .. (855539)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (871619) .. (871619)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1084830) .. (1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1096846) .. (1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1119881) .. (1119881)

OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1130881) .. (1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1310988) .. (1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1313224) .. (1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1349473) .. (1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
  
Alignment Scores:  
Pred. No.: 1.03e-27 Length: 1664976  
Score: 392.50 Matches: 169  
Percent Similarity: 40.1% Conservative: 125  
Best Local Similarity: 23.0% Mismatches: 241  
Query Match: 7.4% Indels: 199  
DB: 3 Gaps: 27  
  
US-09-515-363C-2 (1-1025) x US-09-692-570-1 (1-1664976)  
QY 306 LeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAsnIle 325  
Db 1658788 TTGGAGCGAGGTTGTATGACAGATTATTGACGAATGCTTTAAAGAAAAGACATTTA 1658847  
QY 326 IleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIleAlaLysAsp 345  
Db 1658848 TGTGTT--TTATCGACAGGTTTAGTAAACAGCTATTGCTATTAGTTATAGCAGGT 1658904  
QY 346 HisLeuAspLysLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLys 365  
Db 1658905 ATTTTAACAAAAG-----GATGAAAGGTTTAACTTACCTTCA 1658949  
QY 366 ValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeuLysIleTyr 385  
Db 1658950 AGACCTTTGGTTGAGCAACACTACACAGATTAAACAGGTTTAAACATTGATGAAGAT 1659009  
QY 386 ArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPheProGluValValLys 405  
Db 1659010 AAAATAATAGCTTTAACTGAAATAATCCAGCCAAA--AAGAGCTGAACCTATATAA 1659066  
QY 406 SerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGlu 425  
Db 1659067 AAAGGAAATCTTTATAGCTACACCAAGATTATAGAAAACGATATACATA----- 1659117  
QY 426 AsnGlyGluAspAlaGly---ValGlnLeuSerAspPheSerLeuIleIleAspGlu 444  
Db 1659118 -----GCTGAAGAATAAATGTGATGAATTATTATTATGATGCTGATGAA 1659165  
QY 445 CysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGln 464  
Db 1659166 GCCCACCACACACAGAGACCATGCTAT-----GCATTGTGACCAAAA 1659210  
QY 465 LysLeuLysAsnAsnArgLeuLysGluAsnLysProValIleProLeuProGlnIle 484  
Db 1659211 AAATTTAAGATAAA-----TGTCTATATT 1659234  
QY 485 LeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAlaLysAlaGluGlu 504  
Db 1659235 TTAGGTTTAACGGCATCTCCAGT-----TCTGATATTGAT 1659270  
QY 505 HisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThrValLysGluAsn 524  
Db 1659271 AAAGTCATGAAATTTGTGAAAATTGAATTGACGACGTTGAGTGAAGTGAAGTGAAGAT 1659330  
QY 525 LeuAspGlnLeuLys----- 529

Db 1659331 GATGAGGATGTAACCATACATCTGCTAAAGTAAACTTATCCCAATTAGCAATTGATTTA 1659390  
Qy 530 ---AsnGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThrArgGlu--- 547  
Db 1659391 CCCAACGAATTTAAAGAGCGCTTAAAA-----TTAATAAATGAAGCTTTAAAGAGAGA 1659444  
Qy 548 -----AspProPheLysGluLys 553  
Db 1659445 TTTAAAAATATTAAAAAGATGCTGAGCTTATAATTCCATTCGCCGATGTACAAAAACAGAA 1659504  
Qy 554 LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe 573  
Db 1659505 CTTATTGAGCTAAATAATAGCTATTTCCTAT----- 1659537  
Qy 574 GlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysAlaAlaLysLysGly 593  
Db 1659537 ----- 1659537  
Qy 594 AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle 613  
Db 1659538 -----GATGAAGAAGTGAAGTATGAACCTATATAAAGTTTGTTCAGAGGCTTTAAAACTT 1659591  
Qy 614 AsnAspThrIleArgMetIleAspAla-----TyrThrHisLeuGlu 627  
Db 1659592 ATGCATGCCAAAGAACTCTTAGAGAGTCACAGAAAGAGTGTATTTTAACTATATATAAT 1659651  
Qy 628 ThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspAspSerAspGlu 647  
Db 1659652 AAAATTATCCATGCAAAAGAACAAATCAGCTAAATCTATGTGTTAATGATGAA----- 1659702  
Qy 648 GlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeuLysProLeu 667  
Db 1659702 ----- 1659702  
Qy 668 LysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeu 687  
Db 1659703 AAAGTTAAGAGAGCGCAGTTAATTATTAATGAATCAGATGTAGAACATCCAAATTA--- 1659759  
Qy 688 LysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThr 707  
Db 1659760 -----GGTAAAGTTGTTGATATGTTAAAAAT 1659786  
Qy 708 IleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThr 727  
Db 1659787 ATTTTGGAAAA-----AATAAGATGAGAGAAATTATTATCTTGTCTCAATAT 1659834  
Qy 728 ArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluVal 747  
Db 1659835 AGGGAACACTGTAGAGAGATTGTTAATCTCTTAACCTCAAAAT----- 1659876  
Qy 748 GlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysProMetThr 767  
Db 1659877 GGAATTAAAGCAATAGATTATTAGACAGGCA--AATAAAGAAAGAAAGGGAATGAGT 1659933  
Qy 768 GlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIle 787  
Db 1659934 CAGAAAGACCAATAGAAAGCTTAGAGAGATTAAAAA---GAGGAAGTGTTTTAACTT 1659990  
Qy 788 AlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyr 807  
Db 1659991 TCAACACAGCGTTTCTGAGGAGGAAATAGATATTCATCGGTAATATACATCATATTTTAT 1660050  
Qy 808 GlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGlu 827  
Db 1660051 GAACCACTGCCATCAGAAATTAGGTTTATTTCAGAGGAGAGGTAAGCGATGAGGGGAGAA 1660110  
Qy 828 -----SerThrTyrValLeuVal-----AlaHisSerGlySer 838  
Db 1660111 GGAGGGAAGGTTTATGTTTAATAGCTAAGGGAACAGCTGATGAAGCTTATTAACAGAGAT 1660170  
Qy 839 GlyValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIle 858

Db 1660171 GCTTTATACAAGAAAGGAGATGAAGAGATTATTAAAAATATGTTATTGCTAAAT 1660230  
Qy 859 HisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMet 878  
Db 1660231 AAGAGGTTACAGAAG---AAATTGAAGAAAAAATCTAAAGAGGAATTAAGAGAGACA 1660287  
Qy 879 GlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLys--- 897  
Db 1660288 GAAGAAATAAAAAGAAAAATGAATCAAAAACTGCAGTAAAGAGAAACTTAAGGAG 1660347  
Qy 898 -----AsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerVal 912  
Db 1660348 GAAGAAAGAAAAACCAAAAAAGCCAGTAACGATATTAGATTTCATTAAACAGATTGAAGTT 1660407  
Qy 913 -----LeuAlaCysSerGlyGluAsp----- 919  
Db 1660408 AAGGAAAGTCTAAATCAGAAAGAGATTAATAAACAAGAGATAAAAATTCGGAATAAG 1660467  
Qy 920 ---IleHisValIleGluLysMetHisHisValAsnMetThrPro----- 933  
Db 1660468 CCAATTAAGATTATTGTAGATGTTAGAGAGAGAAATATGGCTTAAGCTTTTACATAATTAT 1660527  
Qy 934 -----GluPheLysGluLeu-----TyrIleValArgGluAsnLysAla 946  
Db 1660528 GCAAAATATTGAGCTAAACACATTAGAAGTGGAGATTATGTTTAACTGATAGGAGTAGTT 1660587  
Qy 947 LeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960  
Db 1660588 GTTGAGAGAAAGACAGCTGAAGACTTGTAAATTCAATAATT 1660629

RESULT 9

US-09-853-768-3  
; Sequence 3, Application US/09853768  
; Patent No. 6444466  
; GENERAL INFORMATION:  
; APPLICANT: Donna T. Ward  
; APPLICANT: Andrew T. Watt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF HELICASE-MOI EXPRESSION  
; FILE REFERENCE: RTS-0217  
; CURRENT APPLICATION NUMBER: US/09/853,768  
; CURRENT FILING DATE: 2001-05-10  
; NUMBER OF SEQ ID NOS: 91  
; SEQ ID NO 3  
; LENGTH: 7037  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (183) ... (5957)  
US-09-853-768-3

Alignment Scores:

Pred. No.:	1.71e-25	Length:	7037
Score:	335.50	Matches:	186
Percent Similarity:	36.5%	Conservative:	113
Best Local Similarity:	22.7%	Mismatches:	233
Query Match:	6.3%	Indels:	288
DB:	3	Gaps:	43

US-09-515-363C-2 (1-1025) x US-09-853-768-3 (1-7037)

Qy 309 ArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAsnIleIleIleCys 328  
Db 315 AGAAATATCAGGTTGAACCTCTTGAAGCAGCTCTGGAT---CATATACCATCGTCTGT 371  
Qy 329 LeuProThrGlySerGlyLysThrArgValAla---ValTyrIleAlaLysAsp----- 345  
Db 372 TTTAAACACTGGCTCAGGGAAGACATTATTGTCTACTACTCTACTTAAGAGCTGTCTC 431  
Qy 346 HisLeuAspLysLysLysLysAlaSerGluProGlyLys---ValIleValLeuValAsn 364  
Db 432 TATCTAGATCTAGGGGAGACTTACGCTAGAAATGGAAGAAAGACGGTCTTGTCAAC 491

QY 365 LysVal---LeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeuLysLys 383  
Db 492 TCTGCAAAACGAGGTGCTCAACAAGTGTACAGCTGTCAAGACTCATTCAGATCTCAAG--- 548  
QY 384 TrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe----- 400  
Db 549 -----GTTGGGGAATACTCAAACCTAGAAAGTAATGCATCTTGGACAAAAGAG 596  
QY 401 -----ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeu 417  
Db 597 AGATGGAAACCAAGAGTTTACTAAGCAC---CAGGTTCATTATGACTGTGCTATGTGCC 653  
QY 418 GluAsnSerLeuLeuAsnLeuGluAsnGlyLysAspAlaGlyValGlnLeuSerAspPhe 437  
Db 654 TTGAATGTTTGA AAAAT-----GTTACTTATCACTGTCAGACATT 695  
QY 438 SerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIle 457  
Db 696 AACCTTTGGTGTTGATGAGTGCATCTTGCATCCCTAGACCAACCCCTATCGAATTT 755  
QY 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysPro 477  
Db 756 ATGAAG-----CTCTGTGA----- 770  
QY 478 ValIlePro---LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAla 496  
Db 771 ATTGTCCATCATGTCCTCGCATTTGGGACTTAACCTGCTTCATT--TTAAATGGGAAA 827  
QY 497 ThrLysGlnAlaLysAlaGluGlnHisIleLeuLysLeuCysAlaAsnLeuAspAlaPhe 516  
Db 828 TGGATCCAGAGATTGGAGAAAAGTTTCAGAACTA----- 866  
QY 517 ThrIleLysThrValLysGluAsnLeuAspGln-----LeuLysAsn 530  
Db 867 ---GAGAAAATTTCTTAAGAGTAATGCTGAACCTGCACTGACTGCTTGAACAGC 923  
QY 531 GlnIleGlnGluProCysLysLysPheAlaIleAlaAsp-----AlaThrArg 546  
Db 924 TATACTTCTCAGCCATGTGAG-----ATTGTGTGATTTGTGACCATTACTGACAGA 977  
QY 547 GluAspProPheLysGluLysLeuLeuGluIle----- 557  
Db 978 AGTGGCTTTATGAAGACTGCTGATGGAATTAGAAGACCACTTAATTTTATCAATGAT 1037  
QY 558 -----MetThrArgIleGlnThrTyrCysGlnMetSerPromet 570  
Db 1038 TGTAAATATATCTGTACATTCAAAAGAAAGAGATTCTACTTTAATTCGAAACAGATACTA 1097  
QY 571 SerAsp-----PheGlyThrGlnProTyrGluGlnTrpAlaIle 583  
Db 1098 TCAGACTGTGCTGCCGTATGTTAGTCTGTGGACCCCTGGTGTGCAGATAAAGTACTGGA 1157  
QY 584 GlnMetGluLysLysAlaAlaLys-----LysGlyAsnArgLysGlu----- 597  
Db 1158 ATGATGTGAAGAACTACAGAAATACATCAACATGAGCAAGAGAGCTGCACAGGAAA 1217  
QY 598 -----ArgValCysAlaGlnHisLeu 604  
Db 1218 TTTTATTTGTTTACAGACACTTTCSTAAGAAAATACATGCATATGTGAAGACACTTC 1277  
QY 605 ArgLysTyrAsn-----GluAlaLeuGlnIleAsnAspThr 616  
Db 1278 TCACCTGCCTCACTTGACCTGAATTTGTAACTCCTAAAGTAATCAAACTGCTCGAAATC 1337  
QY 617 IleArgMetIleAspAlaTyrThrHis-----LeuGluThrPheTyrAsn 631  
Db 1338 TTACGCAAAATATAAACCATATGAGCGACACAGTTTGAAGCGTTGAGTGTATATAAT 1397  
QY 632 GluGluLysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyAlaAspAsp 651  
Db 1398 AGAAATCAGGATTAATTATGTGTATGG-----AGTATTTCTGAGGATGATGAT 1445  
QY 652 GluTyrCysAspGlyAspGluAspGluAspAspLeuLysProLeuLysLeuAspGlu 671

Db 1446 GAG-----GATGAAGAAATTGAAGAAAAGAGAGCCA----- 1478  
QY 672 ThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAla 691  
Db 1478 ----- 1478  
QY 692 GluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGln 711  
Db 1479 -----GAGACAAATTTCTCTCTCTCTTTTACCACAACTTTTGTGC----- 1517  
QY 712 TyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAla 731  
Db 1518 -----GGAATTATTTTGTGGAAAGAAAGATACACAGCA 1550  
QY 732 TyrAlaLeuSerGlnTrpIleThrGluAsnGluLys-----PheAla 745  
Db 1551 GTTGTCTTAAACAGATTGATTAAGGAAGCTGGCCAAACAGATCCAGAGCTGGCTTATATC 1610  
QY 746 GluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHis----- 759  
Db 1611 AGTAGCAATTTTCATAGTGCATGTCATGGGAAAGATCAGCTCGCAACACAGATG 1670  
QY 760 SerSerGluProPheLysPrometThrGlnAsnGlnLysGluValIleSerLysPheArg 779  
Db 1671 GAAGCAGAAATTCA-----AAACAGAAAGAGGTACTTAGAAATTTTGA 1715  
QY 780 ThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys 799  
Db 1716 GCACATGAGACCAACCTGCTTATTTGCCAACAGATTTGTAGAAAGGGTGTGATATACCA 1775  
QY 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819  
Db 1776 AAATGCAACTGTGTGTTGTTGATTGTCGCCACAGAAATATCGATCTTATGTTCAATCT 1835  
QY 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839  
Db 1836 AAAGAAAGACCAAGGCCATCTCTTAATTATATATGTTAGCG----- 1880  
QY 840 ValIleGlnHisGluThrValAsnAspPheArgGluLysMet--MetTyrLysAlaIle 858  
Db 1881 -----GATACAGACAAATAAAAGTTTGAAGAAAGACCTTAACCTAACAAAGCTATT 1934  
QY 859 HisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMet 878  
Db 1934 ----- 1934  
QY 879 GlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsn 898  
Db 1935 -----GAAAAGATCTTGAGAAACAG----- 1955  
QY 899 AsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGlu 918  
Db 1956 -----TGTCCAAAGTCGGTTGATACCTGTTAG 1982  
QY 919 ---AspIleHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGlu 937  
Db 1983 ACTGACATTGATCCTGTGCATGATGATGATGACGTT--TTCCSACCA----- 2027  
QY 938 LeuTyrIleValArgGluAsnLysAla----- 946  
Db 2028 ---TATGTGTGAGCGCTGACGATGTTGCCAGAGTCAACAATCAACACGGCCATTGGA 2084  
QY 947 ---LeuGlnLysLysCysAla-AspTyrGlnIle-----As 957  
Db 2085 CACATCAATAGATACTGTGCTAGATTACCAAGTATCCGTTTACTCATCTAGCTCTAA 2144  
QY 957 nglyGlu-----IleIleCysLysCysGly 965  
Db 2145 TGCAGAACCCGAGAGTTGCTGATGATGATGATTTTATTCAACTCTTATCTGCCAAATTAA 2204  
QY 965 yGlnAlaTrp---GlyThrMetMetValHis-----LysGly 976

Db 2205 TCACCTCTTGAGCCTCCATTGTTGTCACCAATGAGCTGTGTACGATTGGCTGAAAGA 2264  
Qy 976 yLeuAspLeuPro-----CysLeuLysIleArgAsnPhenValVal 990  
Db 2265 GTTGTCGCTTCATTGTCGTGAGAAACTGCACAAAAATTGGCGAAGTGCATGATGCCATT 2322  
RESULT 10  
US-09-408-020-65  
; Sequence 65, Application US/09408020  
; Patent No. 6632937  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; APPLICANT: Schleper, Christa  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCOOP.002A  
; CURRENT APPLICATION NUMBER: US/09/408,020  
; CURRENT FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: 60/102,294  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 65  
; LENGTH: 1512  
; TYPE: DNA  
; ORGANISM: Cenarchaeum symbiosum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1512)  
US-09-408-020-65  
Alignment Scores:  
Pred. No.: 1.52e-21 Length: 1512  
Score: 289.50 Matches: 137  
Percent Similarity: 39.4% Conservative: 88  
Best Local Similarity: 24.0% Mismatches: 188  
Query Match: 5.5% Indels: 158  
DB: 3 Gaps: 23  
US-09-515-363C-2 (1-1025) x US-09-408-020-65 (1-1512)  
Qy 303 GluPro---GluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321  
Db 34 GAGCCCGGCGCGCTCGAGAGCGCGCACTACAGGTGGCGCTTGCAGCAGGCCATACGG 93  
Qy 322 G1yLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341  
Db 94 ---GAAAACTGCATAGTGTGCTGCTTACCGGCTCGGCAAGACGGCGCTGCAG 150  
Qy 342 IleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIleVal 361  
Db 151 GTGATCTCCCACTATTGTGACGAGGAGGCGGCTTCTCTTCCGCGCCGACAGGGTG 210  
Qy 362 LeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeu 381  
Db 211 CTGGTAAAC-----CAGCACCGCCAGTTCTG 237  
Qy 382 LysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-----LysIleSer 399  
Db 238 GGC-----AGGGCCCTTACCATATCCGATATTACCTGTGCACAGGCGAGACACC 288  
Qy 400 PheProGluValValLysSerCys-----AspIleIleIleSerThrAlaGlnIleLeu 417  
Db 289 GTCCCGAGGCGCAAAAAGCTTGGGCGCGCAGCGTGATCTGCGCACCCCGAGATAACA 348  
Qy 418 GluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPhe 437  
Db 349 AGAAACGACATAGCGCGC-----GGAATGGTCCCGCTCGAACAGTTC 390  
Qy 438 SerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIle 457  
Db 391 GGCCTGTTGTTGTCGACGAGGCCCAAGGGCGGTGGCGCACTATGCTATTCGCAATA 450

Qy 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysPro 477  
Db 451 GCG-----CGTGCAGTGGGGAGAACTCT--- 474  
Qy 478 ValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThr 497  
Db 475 -----AGAAATGATCGGCATGACTGCG----- 495  
Qy 498 LysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThr 517  
Db 496 -----ACC 498  
Qy 518 IleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLys 537  
Db 499 CTTCCAAGCGAGAGGAGAAACCCGACGAGATATGGCCACTCTTCTCTCAAG----- 552  
Qy 538 LysPheAlaIleAlaAspAlaThrArgGluAspPro----- 549  
Db 553 -----AGCATAGCACAAAGACCGAAGACGACCGGATGTAAAGCCCTACGTGACGAG 606  
Qy 550 -----PheLysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyr 564  
Db 607 ACCGAAACTGAATGATTAAGGTGAGCTGCCCGGAGATGAAGAGATCCAAAAGCTC 666  
Qy 565 CysGlnMetSerPrometSerAspPheGlyThrGlnProTyrGluGlnTrpAlaIleGln 584  
Db 667 CTGAAGATG-----GCCCTCGAC 684  
Qy 585 MetGluLysLysAlaAlaLysLys-----GlyAsnArg-LysGluArg 598  
Db 685 GAAGATATGCGGCGCTCAGAGGTGCGGCTATGATCTCGGCTCGAACAGGTGCTCTCG 744  
Qy 598 gValCysAlaGluHisIleLysLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleAr 618  
Db 745 GCTTGCTCTCCGCTTCCGATGGTCTTCTTAAGCGGCAACAGCGGCGGCAAAACCTTTG 804  
Qy 618 gMetIleAspAlaTyrThrHisLeuGluLysThrPheTyrAsnGluLys---AspLysLys 637  
Db 805 TTTACTGCATACGATACATACATACGCGCTCAACATATTTCAGAGCCACGGGTCACGCCG 864  
Qy 637 sPhe-AlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyA 657  
Db 865 TTTCTAAAGTTCTGCGAGAGAACCGTCAAGAAAAGG----- 902  
Qy 657 sPgluAspGluAspAspLeuLysLys-ProLeuLysLeuAspGluThrAspArgPheLeu 676  
Db 903 -----CGCCGGTGTTCAGAGCTGTTTCAGAGAGACAGA----- 936  
Qy 677 MetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyr 696  
Db 937 -----AACTTACAGGGGCCATGCGCGCGCAAGCGGCGCAGCGCCGCGCATG 987  
Qy 697 GluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGlu 716  
Db 988 GAGCATCCAAAGATACCAAGTTG-----GAA 1014  
Qy 717 GluSerAlaArgLys-----IleIlePheThrLysThrArgGlnSerAla 731  
Db 1015 GAGGCTGTGCGGCGGCAAGGGAAGCGGTGCTTTACAAGCTACAGGACTCTGTG 1074  
Qy 732 TyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAla 751  
Db 1075 GATTTAATA-----CACTCAAGCTGCAGGCTGCCGGGATAAAGCTCG 1116  
Qy 752 HisHisLeuIleGly---AlaGlyHisSerSerGluPheLysPrometThrGlnAsnGlu 770  
Db 1117 GGGATCTCATAGAAAGCGCGGAGAA-----AAGGCGCTCAAGCAGAAAAAA 1164  
Qy 771 GlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThr 790  
Db 1165 CAGGTAGAGACTGTGCGCAAGTTCCGCGACGGGGGATACGACGTGCTGTACTCAAGA 1224  
Qy 791 ValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuVal 810

Db 1225 GTGGCGAGAGAGGCGCTCGACATATCGAGTAACCTTGTGTATTCTATGCAATGTC 1284  
QY 811 ThrAngluIleAlaMetValGlnAlaArgGlyArg---AlaArgAlaAspGlySerThr 829  
Db 1285 CCAAGCTCGATTAGGTATGTGACAGAGGGCGACCGGACAGAGGACGCGGCAAG 1344  
QY 830 TyrValLeuValAlaHisSerGlySer 838  
Db 1345 CTGGTGTACTGATGGCAAGGGGACT 1371  
RESULT 11  
US-09-408-020-1  
Sequence 1, Application US/09408020  
Patent No. 6632937  
GENERAL INFORMATION:  
APPLICANT: Swanson, Ronald V.  
APPLICANT: Feldman, Robert A.  
APPLICANT: Schleper, Christa  
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
FILE REFERENCE: DORP.002A  
CURRENT APPLICATION NUMBER: US/09/408,020  
PRIOR FILING DATE: 1999-09-29  
PRIOR APPLICATION NUMBER: 60/102,294  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 32998  
TYPE: DNA  
ORGANISM: Cenarchaeum symbiosum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (7604) ... (8908)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (8961) ... (9767)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (10545) ... (10922)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (13944) ... (14612)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (18638) ... (20149)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (20554) ... (20955)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (20956) ... (21834)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (25151) ... (26377)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (27535) ... (28002)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (28065) ... (29483)  
US-09-408-020-1  
Alignment Scores:  
Pred. No.: 3.13e-19 Length: 32998  
Score: 289.50 Matches: 137  
Percent Similarity: 39.4% Conservative: 88  
Best Local Similarity: 24.0% Mismatches: 188  
Query Match: 5.5% Indels: 158  
DB: 3 Gaps: 23  
US-09-515-363C-2 (1-1025) x US-09-408-020-1 (1-32998)

QY 303 GluPro---GluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321  
Db 18671 GAGCCCGGCGCCCTCGAGAGCGCGACTACAGGTGGCCTTGCCGACAGGCCATACGG 18730  
QY 322 GlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341  
Db 18731 ---GAAACTGCATAGTGTGCTGCTTACCGGCTCGGCAAGACGCGGTGCCCTGCAG 18787  
QY 342 IleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIleVal 361  
Db 18788 GTGATCTCCACTATTGTGACGAAGGACGGGGGCTCTTCTTCCGCGGCAAGGTTG 18847  
QY 362 LeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeu 381  
Db 18848 CTGGTAAAC-----CAGACCGCAGTTCCTG 18874  
QY 382 LysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-----LysIleSer 399  
Db 18875 GGC-----AGGCGCCTTACCATATCCGATATTACCTGTCAAGCGGACAGACACC 18925  
QY 400 PheProGluValValLysSerCys-----AspIleIleIleSerThrAlaGlnIleLeu 417  
Db 18926 GTCCGAGCGCGAATAAGCTTGGGGCGGACAGCGTATCTGCCACCCCGAGATAACA 18985  
QY 418 GluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPhe 437  
Db 18986 AGAAACGACATAGCGCC-----GGAATGTCCTCCGCTGCAACAGTTC 19027  
QY 438 SerLeuIleIleIleAspGluCysHisThrAsnLysGluAlaValTyrAsnAsnIle 457  
Db 19028 GGCCTGTGTGTTGACGAGGCCACAGGGCGGTGGCGACTATGCTTATCCGCAATA 19087  
QY 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysPro 477  
Db 19088 GCG-----CGTGACGTGGGGAGAACTCT-- 19111  
QY 478 ValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThr 497  
Db 19112 -----AGATGATCGGCATGACTGCG----- 19132  
QY 498 LysGlnAlaLysAlaGluGlnHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThr 517  
Db 19133 -----ACC 19135  
QY 518 IleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnIleProCysLys 537  
Db 19136 CTTCCAGCGAGAGGAGAAAGCCGACGAGATAATGGCACTCTTCTTCAAG----- 19189  
QY 538 LysPheAlaIleAlaAspAlaThrArgGluAspPro----- 549  
Db 19190 -----AGCATAGCACAAAGGACGAAAGACCGGATGTAAAGCCCTACGTGACGAG 19243  
QY 550 -----PheLysGluLysLeuLeuGlnIleMetThrArgIleGlnThrTyr 564  
Db 19244 ACCGAACTGAATGATTAAGGTGAAGCTGCCCGGAGATGAAGAGATCCAAAAGCTC 19303  
QY 565 CysGlnMetSerPrometSerAspPheGlyThrGlnProTyrGluGlnTyrAlaIleGln 584  
Db 19304 CTGAAGATG-----GCCCTCGAC 19321  
QY 585 MetGluLysLysAlaAlaLysLys-----GlyAsnArg-LysGluArg 598  
Db 19322 GAAAGATATGCGGCCCTCAAGAGGTGCGGCTATGATCTCGGCTCGAACAAGTCTCTCG 19381  
QY 598 GValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArg 618  
Db 19382 GCTCTGCTCGCCTTGCATGGTCTGTTCTAAGCGGCAAGCGGCGGCAAGCCTTTG 19441  
QY 618 gMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLys--AspLysLys 637  
Db 19442 TTTACTGCGATACGATACATACGCGCTCAACATATTGAGGCCACAGGGTACGCG 19501  
QY 637 sphe-AlaValIleGluAspAspSerAspGluGlyLysAspAspGluTyrCysAspGlyA 657

Db 19502 TTCTTAAAGTTCGAGAGACCGTCAAGAAAAAGG----- 19539  
QY 657 spGluAspGluAspAspleuLys- ProleuLysLeuAspGluThrAspArgPheLeu 676  
Db 19540 -----CGCCGGTGTTCAGAGCTGTTCAGAGGAGACAGA----- 19573  
QY 677 MetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyr 696  
Db 19574 -----AACTTTACAGGGGCCATGGCGCGCAAGCGCGCAGCGCGGCATG 19624  
QY 697 GluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGlu 716  
Db 19625 GAGCATCCAAAGATACCMAAGTTG-----GAA 19651  
QY 717 GluSerAlaArgGly-----IleIlePheThrLysThrArgGlnSerAla 731  
Db 19652 GAGGCTGTGCGCGGGCCAAAGGAGCGCGTGTCTTACAAGCTACAGGAGCTGTGTC 19711  
QY 732 TyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAla 751  
Db 19712 GATTTAATA-----CACTCAAGCTGCAGGCTGCCGGGATAAACTCG 19753  
QY 752 HisHisLeuIleGly--AlaGlyHisSerSerGlnPheLysProMetThrGlnAsnGlu 770  
Db 19754 GGGATCCTCATAGAAAGCGCGAGAA-----AAGGGCTCAAGCAGAAAAAA 19801  
QY 771 GlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThr 790  
Db 19802 CAGGTAGAGACTGTCCGCAAGTTCGCGACGCGGGAGTACAGACTGTCTGATCTACAAGA 19861  
QY 791 ValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuVal 810  
Db 19862 GTGGGAGAGAGGCGCTCGACATATCGGAGTAACTGTGTATTCATGCAATGTC 19921  
QY 811 ThrAsnGluIleAlaMetValGlnAlaArgGlyArg--AlaArgAlaAspGluSerThr 829  
Db 19922 CCAAGCTCGATTAAGGTATGTGACAGAAAGGGCAGACCGGCAGGAAGACCGGGCAAG 19981  
QY 830 TyrValLeuValAlaHisSerGlySer 838  
Db 19982 CTGGTGTAATGATGGCAAGGGGACT 20008

RESULT 12  
US-09-408-020-33  
; Sequence 33, Application US/09408020  
; Patent No. 6632937  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; APPLICANT: Schleper, Christa  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCorp.002A  
; CURRENT APPLICATION NUMBER: US/09/408,020  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: 60/102,294  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 33  
; LENGTH: 1509  
; TYPE: DNA  
; ORGANISM: Cenarchaeum symbiosum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1509)  
US-09-408-020-33

Alignment Scores:  
Pred. No.: 6.17e-21 Length: 1509  
Score: 284.00 Matches: 133  
Percent Similarity: 41.4% Conservative: 94  
Best Local Similarity: 24.3% Mismatches: 207

Query Match: 5.3% Indels: 114  
DB: 3 Gaps: 23  
US-09-515-363c-2 (1-1025) x US-09-408-020-33 (1-1509)  
QY 303 GluPro---GluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321  
Db 34 GAGCCCGGTCCGTCGAGAGCGCGCACTACCAAGTGGCGCTGGCGGAACAGCCATACGG 93  
QY 322 GlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341  
Db 94 --GAGAACTGTATCGTGTGCTCCGACGGGCTCGGCAAGACTGCCGTGCCCTCCAG 150  
QY 342 IleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIleVal 361  
Db 151 GTGATCGCCCACTATCTGCAGCAGGCGCGCGGCGCTTCTTCCCTGACCAAGGTC 210  
QY 362 LeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeu 381  
Db 211 CTGGTAAAC-----CAGCACCGCCAGTTCCTG 237  
QY 382 LysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-----LysIleSer 399  
Db 238 GGC-----AGGGCCCTTACCATATCCGATATTACACTGTGTACCGGAGAGACACC 288  
QY 400 PheProGluValValLysSerCys-----AspIleIleIleSerThrAlaGlnIleLeu 417  
Db 289 ATTCCCGCGCGCAAAAGCGGTGGGAGCAGCGTATCTGCCACGCCGAGATAGCA 348  
QY 418 GluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPhe 437  
Db 349 AGAATGATATA-----GAGCGCGCGCTGTCCCGCTGAACAGTTTC 390  
QY 438 SerLeuIleIleIleAspGluCysHisIleThrAsnLysGluAlaValTyrAsnAsnIle 457  
Db 391 GGCTGTGCATATTCCAGCAGGCCACAGGGCGGTGGCGACTATGCCATTCTTCATA 450  
QY 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysValGluAsnLysPro 477  
Db 451 GCGCGG-----GCG 459  
QY 478 ValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThr 497  
Db 460 GTAGGGGATACTCCAGAGTGTGGCATGACGACGCTT-----CCCAGC 507  
QY 498 LysGlnAlaLysAlaGluGlnHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThr 517  
Db 508 GAGAGGAGAGAAGCAGACGAG-----ATAATGGCACCCGTCTCTCCAGAGAC 555  
QY 518 IleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnIleProCysLys 537  
Db 556 ATAGCCCAAGAGACAGACAGACCCGACCGTAAGCCCTATGTATACAGAGACTGCCACC 615  
QY 538 LysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIle 557  
Db 616 GAGTGGATAAAGGTGATCTTCCCCCGAGATGAAGAGATACAGAGGCTCTCAAGCTG 675  
QY 558 MetThrArgIleGlnThrTyrCysGlnMetSerProMetSer--AspPheGlyThrGln 576  
Db 676 GCCCTCGAC--GAGAGGTATTCTCCCTCAAGAGGTGCGGTACGATCTTGCTCGAAC 732  
QY 577 ProTyrGluGlnTrpAlaIleGlnMetGluLysAlaAlaLysLysGlyAsnArgLys 596  
Db 733 AGGTGCTCTCGGCGTCTCCGGCTG--CGCATGGTGTGTCTTGCGGCAACAGGCGC 789  
QY 597 GluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThr 616  
Db 790 GCG-----GCCAAGCCGCTGTCACTGCGATACGCAATAACGTACGCG 831  
QY 617 IleArgMetIleAspAlaTyr-----ThrHisLeuGluLutThrPheTyrAsnGluLys 634  
Db 832 CTAAACATATTTCAGAGCGCAGCGGGTCAACGCCCTTTCTAAAGTTCTGCGAGAGACCTCC 891

QY 635 AsplylysphealavalilegluaspaSeraSpgluglylaSpaSpglutyrCys 654  
Db 892 AAGAAAAAGGCGCTCGCGCGCGAG----- 918  
QY 655 AspGlyAspGluAspGluAspLeuLysProleuLysLeuAspGluThrAspArg 674  
Db 919 -----CTGTTCAAGAGGACCGG 936  
QY 675 pheuMetThrleuPheGluAsnAnLysMetleuLysArgLeuAlaGluAsnPro 694  
Db 937 -----AACTTTACAGGGCCATCGCGCGCAAGGCCGCGAGCGGCA 981  
QY 695 GluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrileMetGluGlnTyrThrArg 714  
Db 982 GGCATGAGCATCCCAAGATACCAAGCTCGAGGATGCCGTC----- 1023  
QY 715 ThrGluGluSerAlaArgGly-----IleIlePheThrLysThrArgGlnSerAlaTyr 732  
Db 1024 -----CGCGGGGCCCGGGGAAGCGCGCTGCTTTACGAGCTATCGTGAATTCTGCGAC 1077  
QY 733 AlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyVallyAlaHis 752  
Db 1078 CTCATA-----CACTCAAGACTCAAGCGCGCGGGATTAACCTCGGGC 1119  
QY 753 Hisleuilegly---AlaGlyHisSerSerGluPheLysPromeThrGlnAsnGluGln 771  
Db 1120 ATCTGATAGGAAAGCGGAGAA-----AAGGGCCTAAAGCAGAGAAACAG 1167  
QY 772 LysGluValIleSerLysPheArgThrGlyLysIleAsnLeuIleAlaThrThrVal 791  
Db 1168 GTGAGACTGTGGCAAGTTCGTCGACGCGCGGTACAGCTGTGTATCGACGAGGCTC 1227  
QY 792 AlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThr 811  
Db 1228 GGGCAGAGGGGCTCGACATATCGAGGTCAACCTGTGATATTCTATGACATGTGCCA 1287  
QY 812 AsnGluIleAlaMetValGlnAlaArgGlyArg---AlaArgAlaAspGluSerThrTyr 830  
Db 1288 AGCTGCATCAGGTACGTGACAGAGAGGGGAGAGACAGCAGAAAGACCGCGCAGGCTG 1347  
QY 831 ValLeuValAlaHisSerGlySer 838  
Db 1348 ATAGTATTGATGGCAAGGGGACG 1371  
RESULT 13  
US-09-408-020-2  
; Sequence 2, Application US/09408020  
; Patent No. 6632937  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; APPLICANT: Schleper, Christa  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCOPT.002A  
; CURRENT APPLICATION NUMBER: US/09/408,020  
; CURRENT FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: 60/102,294  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 42432  
; TYPE: DNA  
; ORGANISM: Cenarchaeum symbiosum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3)...(10421)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (10625)...(11434)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (11478)...(13046)

; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (13046)...(14620)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (23558)...(24862)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (24913)...(25728)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (26504)...(26881)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (29655)...(30491)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (34559)...(36067)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (37002)...(37403)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (37404)...(38282)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (39454)...(40572)  
US-09-408-020-2  
Alignment Scores:  
Pred. No.: 1.97e-18 Length: 42432  
Score: 284.00 Matches: 133  
Percent Similarity: 41.4% Conservative: 94  
Best Local Similarity: 24.3% Mismatches: 207  
Query Match: 5.3% Indels: 114  
DB: 3 Gaps: 23  
US-09-515-363C-2 (1-1025) x US-09-408-020-2 (1-42432)  
QY 303 GluPro---GluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321  
Db 34592 GAGCCCGGTGCCGTGAGAGCGCGACTACAGGTGGCCTGCGCGAAGCAGCCATACGG 34651  
QY 322 GlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341  
Db 34652 ---GAGAACTGTATCGGTGCTCCCGACGGGCTCGGCAAGACTGCCCTCCAG 34708  
QY 342 IleAlaLysAspHisIleAspLysLysLysAlaSerGluProGlyLysValIleVal 361  
Db 34709 GTGATCGCCCACTATCTGACGAGGGCGCGGCGCTCTTCCTTGCCCTCAAGGGTC 34768  
QY 362 LeuValAsnLysValLeuLeuValGluGlnPheArgLysGluPheGlnProPheLeu 381  
Db 34769 CTGGTAAAC-----CAGCACCGCAGTTCTTG 34795  
QY 382 LysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-----LysIleSer 399  
Db 34796 GGC-----AGGGCCCTTACCATATCCGATATTAACACTGTCACGGAGAGACACC 34846  
QY 400 PheProGluValValLysSerCys-----AspIleIleIleSerThrAlaGlnIleLeu 417  
Db 34847 ATTCCTCCGGCGCAAAAGCGGTGGGAGGCGAGCTGATTCGCCGACGCCGAGATAGCA 34906  
QY 418 GluAsnSerLeuLeuAsnLeuGluAsnGlyLysAspAlaGlyValGlnLeuSerAspPhe 437  
Db 34907 AGAAATGATATA-----GAGCGCGGCGCTGATCCCGCTCGAACAAGTTC 34948  
QY 438 SerLeuIleIleIleAspGluCysHisIleThrAsnLysGluAlaValTyrAsnAsnIle 457  
Db 34949 GGCTGTGATATTCGACGAGGCCACAGAGCGGTGGCGACTATAGCTTATTCCTTCATA 35008  
QY 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysPro 477

Db 35009 GCGCG-----GCG 35017  
Qy 478 ValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThr 497  
Db 35018 GTAGGGATAACTCCAGAGTGGTCATGACTGCAGCGCTT-----CCGAGC 35065  
Qy 498 LysGlnAlaLysAlaGluGlnHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThr 517  
Db 35066 GAGAGGAGAAGGACAGACGAG-----ATAATGGGCACCCTGCTCTCCAGAGC 35113  
Qy 518 IleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLys 537  
Db 35114 ATAGCCCAAGACAGACAGACCGGACCGCTAAAGCCCTATGTACAGAGACTGCCACC 35173  
Qy 538 LysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuGluIle 557  
Db 35174 GAGTGGATAAAGGTGATCTTCCCCCGAGATGAAGAGATACAGAGGCTCTCAAGCTG 35233  
Qy 558 MetThrArgIleGlnThrTyrCysGlnMetSerProMetSer--AspPheGlyThrGln 576  
Db 35234 GCCCTCGAC--GAGAGGTATTCCTCCCTCAAGAGGTGCGGCTACGATCTTGCTCGAAC 35290  
Qy 577 ProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysGlyAsnArgLys 596  
Db 35291 AGGTGCTCTCGCGCTGCTCGGCTG--CGCATGTGTGCTTGCGCGCAAGCGGC 35347  
Qy 597 GluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThr 616  
Db 35348 GCG-----GCCAAGCCGCTGTTCACTGCATACGATACGATACGATACGCG 35389  
Qy 617 IleArgMetIleAspAlaTyr-----ThrHisLeuGluThrPheTyrAsnGluLys 634  
Db 35390 CTAAACATATTCGAGCGCACGGGGTCACGCCCTTTCTAAAGTTCTGCGAGAGACCTCC 35449  
Qy 635 AspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCys 654  
Db 35450 AAGAAAGGCGCTCGCGCGGAG----- 35476  
Qy 655 AspGlyAspGluAspGluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArg 674  
Db 35477 -----CTGTTCAACAGAACCGG 35494  
Qy 675 PheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnPro 694  
Db 35495 -----AACTTTACAGGGCCATCGCGCCGCAAAAGCGCGCACGCGGCA 35539  
Qy 695 GluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArg 714  
Db 35540 GGCATGAGCATCCCAAGATACAAAGCTCGAGATCGCTC----- 35581  
Qy 715 ThrGluLysSerAlaArgGly-----IleIlePheThrLysThrArgGlnSerAlaTyr 732  
Db 35582 -----CGCGGGCGCGGAAAGCGCTGTTCTTACGAGCTATCGTATGTTGTCGAC 35635  
Qy 733 AlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAlaHis 752  
Db 35636 CTGATA-----CACTCAAGACTCAAGCGCGCGGATAAACTCGGGC 35677  
Qy 753 HisLeuIleGly--AlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluGln 771  
Db 35678 ATCTGTATAGCAAAAGCGGGAGAA-----AAGGGCCTAAAGCAGAGAAACAG 35725  
Qy 772 LysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrVal 791  
Db 35726 GTGAGACTGTGCAAAAGTTCGTACGCGCGGTACGACGTGCTGTATCGACGAGGCTC 35785  
Qy 792 AlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThr 811  
Db 35786 GCGGAGAGGGGCTGCACATATCGAGGTCAACCTGTGATATTCTATGACATGTGCCA 35845  
Qy 812 AsnGluIleAlaMetValGlnAlaArgGlyArg--AlaArgAlaAspGluSerThrTyr 830  
Db 35846 AGCTCATTCAGGTACGTGCAGAGAGGGGAGAAACAGCAGAAAGACGCGCGCAGGCTG 35905

Qy 831 ValLeuValAlaHisSerGlySer 838  
Db 35906 ATAGTATGTATGCAAGGGGACG 35929  
RESULT 14  
US-10-104-047-1641  
; Sequence 1641, Application US/10104047  
; Patent No. 6943241  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. 6943241e1 full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; PRIOR FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1641  
; LENGTH: 2354  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-104-047-1641  
Alignment Scores:  
Pred. No.: 1.51e-20 Length: 2354  
Score: 283.50 Matches: 166  
Percent Similarity: 35.2% Conservative: 116  
Best Local Similarity: 20.7% Mismatches: 245  
Query Match: 5.3% Indels: 277  
DB: 3 Gaps: 35  
US-09-515-363C-2 (1-1025) x US-10-104-047-1641 (1-2354)  
Qy 235 LysGluValTrpGlyMetGluAsnAsnSerSerGluSerSerPheAlaAspSer----- 252  
Db 6 AAGGAAACCGATGGGATCGGAACCGTAGCCGTTGAGCTGCTGCTGTACGATATCTGA 65  
Qy 253 -----SerValValSerGluSerAspThrSerLeuAlaGluGlySerValSerCys 269  
Db 66 CAGAACCTTCGGTGGTT-----GTGGCCTTAATGACGGCAGCAAGAACG 113  
Qy 270 LeuAspGluSerLeuGlyHisAsnSerAsnMetGlySerAspSerGlyThrMetGlySer 289  
Db 114 CTTTTCAGACGTGGGC-----TCAAGTATTCGCCGATCATCTGGACTCCGGGTGC 167  
Qy 290 AspSer----- 291  
Db 168 AGCTCCGGAAGTGAAGCACTCAGAGCCCTGGAGCTCCAAGCGCCTTTGCCAGACGA 227  
Qy 292 -----AspGluGluAsnValAlaAlaArgAlaSerProGluProGlu 305  
Db 228 GCGGAGGCTCAGCTGAGTGGACGATGATGTGTTGCTGTGCGGCGGTACGAGGCTGAG 287  
Qy 306 LeuGln----- 307  
Db 288 CCGCAGTTCTGTCTAGAAATGGCGGTTCTGCACCTCCGCGGCGCCTGTGATTTAC 347  
Qy 308 -----LeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGly 322  
Db 348 CTTACCAATTGCCAGTGGGGAAGTACACAGCTGCACATTTCGCGGCTGCTG--TTT 404  
Qy 323 LysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIle 342  
Db 405 TGCATACGCTGATGTCTGCCTTACCGGACTGGGAAAGACCTTATATGCGCGGTGTC 464  
Qy 343 AlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIleValLeu 362  
Db 465 ATGTACATTTCTAC-----CGCTGTTCCCTTCAGGAAAGGTGCTTCATG 512  
Qy 363 Val---AsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeu 381

```
Db      513 GCCCCAACGAACCTTGGTGACACAGCAG-----ATC 545
Qy      382 LysLysTrpTyrArgValIleGly-----LeuSerGlyAsp 393
      ::::|||||:
Db      546 GAGGCTTGCTACCAGGTGATGGGTATCCCGCAATCCCATGCGCCGAATGACAGGGTCT 605
Qy      394 ThrGlnLeuLysIleSerPheProGluValValYsSerCysAspIleIleIleSerThr 413
      |||||:
      ::::|||||:
Db      606 ACACAAAGCTTCC--ACCAAGAAAGAAATATGGTCAGTAAGAGATGCTTTTCTTACA 662
Qy      414 AlaGlnIleLeuGluAsnSerLeuLeuAsnGluAsnGlyGluAspAlaGlyValGln 433
      |||||:
      ::::|||||:
Db      663 CCTCAGGTCAATGTAATGACCTT-----TCTAGAGAGCTTGTCCC 704
Qy      434 LeuSerAspPheSerLeuIleIleLeuAspGluCysHisIleThrAsnLysGluAlaVal 453
      ::::|||||:
      ::::|||||:
Db      705 GCTGCTGAATAAAGTGTATGATTATGATGAGCTCAATAAGCTCTCGAAACTATGCT 764
Qy      454 TyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLys 473
      |||||:
      ::::|||||:
Db      765 TATTGCCAGGCTGTAAAGAA-----CTAGTCAATATACAAATCACCCTT----- 809
Qy      474 GluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGly--- 492
      ::::|||||:
      ::::|||||:
Db      810 -----AGAACTTGGCTCTAAGTGCCACACACAGGTAGT 842
Qy      493 -----ValGlyGlyAla----- 496
      ::::|||||:
Db      843 GATATAAAGGCTGTGCAACAAGTTATTACTAACCTGCTAATTGGCAGATAGAGCTTCGT 902
Qy      497 -----ThrLysGlnAlaLysAlaGluGlnHisIle 506
      ::::|||||:
      ::::|||||:
Db      903 TCTGAAGATTCTCCAGATATTTGACATATTCTCATGAAGAAAGTTGAAAGCTTATT 962
Qy      507 LeuLysLeuCysAlaAsn-----LeuAspAla 515
      ::::|||||:
      ::::|||||:
Db      963 GTTCGGCTTGGTGAAGACCTTGCAAGCCATCCAAAGACCTATATCCAGATTGGAAATCA 1022
Qy      516 PheThrIleLysThrValLysGluAsnLeu----- 525
      |||||:
      ::::|||||:
Db      1023 TTTCGCTGTTCTTTGATTCAAGGAATGTTTGGATGAGAGGATATCCCAATCTAACA 1082
Qy      526 -----AspGlnLeu--LysAsnGlnIleGlnLysProCys 536
      ::::|||||:
      ::::|||||:
Db      1083 AAATATCAGATAATTCTGGCAAGATCAGTTTAGAAAAACCC-ATCTCCGAATATTGT 1141
Qy      537 LysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGlyLysLeuGlu 556
      ::::|||||:
      ::::|||||:
Db      1142 -----GGGAATACAAACAAGCATAAATCG---AGGAGAGTTTGCTATT 1182
Qy      557 IleMetThr-----ArgIleGlnThrTyrCysGlnMetSerPrometSerAspPheGly 574
      ::::|||||:
      ::::|||||:
Db      1183 GATATTAGTTTATATCATGGTTATGAATTATTGACAGCAAAATGGGAATGAGATCATTAATT 1242
Qy      575 ThrGlnProTyrGluGlnThrAlaIleGlnMetGlyLysAlaAlaLysLysGlyAsn 594
      ::::|||||:
      ::::|||||:
Db      1243 TCTTCCTTGTGGAATTATGG-----ATGGAAT-AAAAGGATG----- 1280
Qy      595 ArgLysGluArgValCysAlaGlnHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsn 614
      ::::|||||:
      ::::|||||:
Db      1281 -----ACACGGTCAAAAAATGAACCTTGCGCCGAAATGAA 1313
Qy      615 AspThrIleArgMetIleAspAlaTyrThrHisLeuGlnThrPheTyrAsnGluGlyLys 634
      |||||:
      ::::|||||:
Db      1314 GACTTCATGAAGACTC-----TATAATCATCTAGAGTGTATGTTTGACAGCTACACGT 1364
Qy      635 AspLysLysPheAlaValIleGluAspAspSerAspGluGlyLysAspAspGluTyrCys 654
      ::::|||||:
      ::::|||||:
Db      1365 -----AGTACTTCAGCAAAATGCTATTCTGCTATCCAA 1397
Qy      655 AspGlyAspGluAspGluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArg 674
      |||||:
      ::::|||||:
Db      1398 CAAGGAGAT----- 1406
```

```
Qy      675 PheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnPro 694
Db      1407 -----AAAAATAAAAATTTGTT----- 1424
      ::::|||||:
Qy      695 GluTyrGluAsnGluLysLeuThrLysLysLeuArgAsnThrIleMetGluGlnTyr----- 712
      ::::|||||:
      ::::|||||:
Db      1425 ---TATAGTCATCCAAAGTTAAAGAAATTGAGAAAGTTGTAATTGAACTTCAAGTCA 1481
Qy      713 -----ThrArgThrGluGluSerAlaAArgLysIleIlePheThr 725
      ::::|||||:
      ::::|||||:
Db      1482 TGGAAATGCTGAATAACACTACTGAAGAAGAACGTGATGAGACCCGAGTTATGATCTTCTCT 1541
Qy      726 LysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAla 745
      ::::|||||:
      ::::|||||:
Db      1542 TCATTTCGAGATAGTGTTCAGAAATTCAGAAATGCTTTCACAGCATCAGCCAAATTATT 1601
Qy      746 GluValGlyValLysAlaHisIleLysLeuIleGlyAlaGlyHisSerSer-----GluPhe 763
      ::::|||||:
      ::::|||||:
Db      1602 AGAGTA-----ATGACTTTTGTGCGGCATGCTCAGGGAAGCAGC 1643
Qy      764 LysPrometThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIle 783
      ::::|||||:
      ::::|||||:
Db      1644 AAGGTTTTCACCAAGAGCAACTGAGGTAGTGAACAGATTTCGACGCGTGTAC 1703
Qy      784 AsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlyCysAsnIle 803
      ::::|||||:
      ::::|||||:
Db      1704 AACACGCTGTTTCTACTGTGTGGGTGAAGAGTTTGGATATAGAGAAAGTTGATCTT 1763
Qy      804 ValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArg--- 822
      ::::|||||:
      ::::|||||:
Db      1764 ATATATGTTTGTGATTTCCCAAGAGAGCCCAATTCGCTTGTACACAGATGGTAGAACT 1823
Qy      823 AlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGlu 842
      ::::|||||:
      ::::|||||:
Db      1824 GGCCGTAAACGTCAAGGAGGATAGTTATTATCTTCTTGAAGGA----- 1868
Qy      843 HisGluThrValAsnAspPheArgGlyLysMetMetTyrLysAlaIleHisCysValGln 862
      ::::|||||:
      ::::|||||:
Db      1869 -----CGAGAGGAACGTATTATAATCAG-----AGTCAG 1898
Qy      863 AsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMet 882
      ::::|||||:
      ::::|||||:
Db      1899 TCCACAAAAAGAGTATATAT----- 1919
Qy      883 GluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeu 902
      ::::|||||:
      ::::|||||:
Db      1920 ---AAAGTATTTCAGTAACAGCAGGCTCTTCATTTTTCACCAAGAGTCCACGAATG 1976
Qy      903 Ile 903
      ::::
Db      1977 GTT 1979

RESULT 15
US-08-143-576-6
; Sequence 6, Application US/08143576
; Patent No. 5643761
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: METHOD FOR GENERATING A SUBTRACTED
; TITLE OF INVENTION: CDNA LIBRARY AND USES OF THE GENERATED LIBRARY
; CORRESPONDENCE ADDRESS:
; ADDRESS: 30 Rockefeller Plaza
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/143,576  
FILING DATE: 25-OCT-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 43563/JPW/AKC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523 COOP UI  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 193 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-143-576-6

Alignment Scores:  
Pred. No.: 3.28e-16 Length: 193  
Score: 227.50 Matches: 50  
Percent Similarity: 75.4% Conservative: 2  
Best Local Similarity: 72.5% Mismatches: 10  
Query Match: 4.3% Indels: 8  
DB: 2 Gaps: 1

US-09-515-363C-2 (1-1025) x US-08-143-576-6 (1-193)

QY	947	LeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlnIleIleCysLysCysGlyGln	966
DB	1	CTGCAAAAGAAGTGTGCCGACTATAA-ATAAATGGTGAATCATCTGCAAAATGTGCCAG	59
QY	967	AlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLysIleArgAsn	986
DB	60	GCTTGGGGAAACAATGATGTGCACAAAGCCTTAGATTGCTGTCTCAAAATTAAGGAAT	119
QY	987	PheValValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTrpValGluLeu	1006
DB	120	TTTGTAGTGGTTCAAAA-----TATCACAAAGACGTACAAAGTG	158
QY	1007	ProIleThrPheProAsnLeuAspTyr	1015
DB	159	GTAGATACTATCACATTCACACTGACTAT	185

Search completed: March 12, 2006, 06:36:44  
Job time : 5367 secs

**This Page Blank (uspto)**